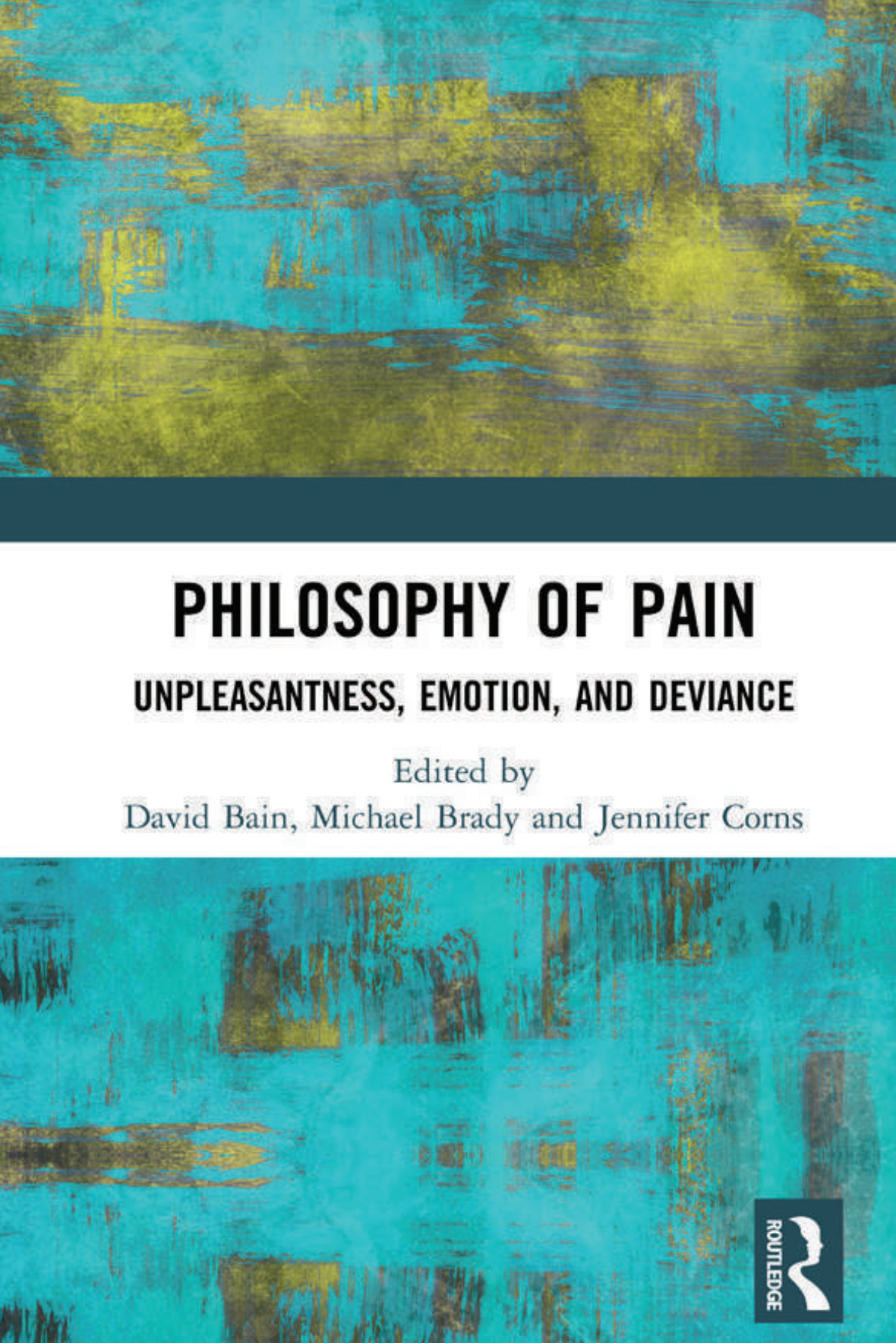




PHILOSOPHY OF PAIN

UNPLEASANTNESS, EMOTION, AND DEVIANCE

Edited by

David Bain, Michael Brady and Jennifer Corns

ROUTLEDGE



Philosophy of Pain

Over recent decades, pain has received increasing attention as philosophers, psychologists, and neuroscientists try to answer deep and difficult questions about it. What is pain? What makes pain unpleasant? How is pain related to the emotions? This volume provides a rich and wide-ranging exploration of these questions and important new insights into the philosophy of pain. Divided into three clear sections – pain and motivation, pain and emotion, and deviant pain – the collection covers fundamental topics in the philosophy and psychology of pain. These include pain and sensory affect, the neuroscience of pain, pain and rationality, placebos, and pain and consciousness.

Philosophy of Pain: Unpleasantness, Emotion, and Deviance is essential reading for students and researchers in philosophy of mind, philosophy of psychology, cognitive and behavioural psychology, as well as those in health and medicine researching conceptual issues in pain.

David Bain is Reader in Philosophy at the University of Glasgow, UK.

Michael Brady is Professor of Philosophy at the University of Glasgow, UK.

Jennifer Corns is Lecturer in Philosophy at the University of Glasgow, UK.

Philosophy of Pain

Unpleasantness, Emotion, and Deviance

Edited by David Bain, Michael Brady,
and Jennifer Corns

First published 2019
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

and by Routledge
711 Third Avenue, New York, NY 10017

Routledge is an imprint of the Taylor & Francis Group, an informa business

© 2019 selection and editorial matter, David Bain, Michael Brady, and
Jennifer Corns; individual chapters, the contributors

The right of David Bain, Michael Brady, and Jennifer Corns to be identified
as the authors of the editorial material, and of the authors for their individual
chapters, has been asserted in accordance with sections 77 and 78 of the
Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this book may be reprinted or reproduced or
utilised in any form or by any electronic, mechanical, or other means, now
known or hereafter invented, including photocopying and recording, or in any
information storage or retrieval system, without permission in writing from
the publishers.

Trademark notice: Product or corporate names may be trademarks or registered
trademarks and are used only for identification and explanation without intent
to infringe.

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

Library of Congress Cataloguing-in-Publication Data

Names: Bain, David (Reader in Philosophy), editor.

Title: Philosophy of pain: unpleasantness, emotion, and deviance/edited by
David Bain, Michael Brady, and Jennifer Corns.

Description: 1 [edition]. | New York: Routledge, 2018. |

Includes bibliographical references and index.

Identifiers: LCCN 2018004021 | ISBN 9780815361640 (hardback: alk. paper) |

ISBN 9781351115865 (e-book)

Subjects: LCSH: Pain—Philosophy.

Classification: LCC BJ1409 .P45 2018 | DDC 128/.4—dc23

LC record available at <https://lcn.loc.gov/2018004021>

ISBN: 978-0-8153-6164-0 (hbk)

ISBN: 978-1-351-11586-5 (ebk)

Typeset in Goudy

by Deanta Global Publishing Services, Chennai, India

Contents

<i>List of contributors</i>	vii
Introduction	1
DAVID BAIN, MICHAEL BRADY, AND JENNIFER CORNS	
 PART I	
Pain, unpleasantness, and motivation	11
 1 Imperativism and pain intensity	13
COLIN KLEIN AND MANOLO MARTÍNEZ	
 2 Reasons and theories of sensory affect	27
MURAT AYDEDE AND MATTHEW FULKERSON	
 3 A neuroscience perspective on pleasure and pain	60
DAN-MIKAEL ELLINGSEN, MORTEN L. KRINGELBACH, AND SIRI LEKNES	
 PART II	
Pain and emotion	79
 4 The rationality of emotional and physical suffering	81
MICHAEL BRADY	
 5 The placebo effect	95
JENNIFER CORNS	
 6 What is the affective component of pain?	120
JESSE PRINZ	

PART III

Deviant pain 145

7 The unpleasantness of pain for humans and other animals 147

ADAM SHRIVER

**8 When is a pain not a pain? The challenge of disorders
of consciousness** 163

VALERIE GRAY HARDCASTLE

9 The first-person in pain 178

FRÉDÉRIQUE DE VIGNEMONT

Bibliography 193

Index 222

Contributors

Murat Aydede is Professor of Philosophy at the University of British Columbia, Canada. He works in the areas of philosophy of mind and moral psychology, specializing in theories of phenomenal consciousness, perception, pain, and sensory affect.

David Bain is Reader in the Department of Philosophy at the University of Glasgow, Scotland. He works in philosophy of mind, with a particular focus on phenomenal consciousness, perception, and the affective dimensions of our mental lives. Jointly with Professor Michael Brady, he was Principal Investigator of the Pain Project (2012-13) and the Value of Suffering project (2013–16). He has published on pain, affect, colour, and Wittgenstein's private language argument.

Michael Brady is Professor of Philosophy in the Department of Philosophy at the University of Glasgow, Scotland. His research centres on the philosophy of emotion, and its links with moral philosophy and epistemology. In 2013 his book *Emotional Insight* was published by Oxford University Press. He was recently Principal Investigator, jointly with Dr. David Bain, on a major interdisciplinary project on the Value of Suffering, hosted at Glasgow; his book on this topic, *Suffering and Virtue*, will be published by Oxford University Press in 2018. He was Director of the British Philosophical Association from 2011 to 2014, having previously served as Secretary of the Scots Philosophical Association. Outside of academia, he is philosopher-in-residence for the Manchester-based theatre company Quarantine.

Jennifer Corns is Lecturer in Philosophy at the University of Glasgow, Scotland. Her published research focuses on pain and affect. She aims to use philosophical tools and evaluate empirical research to make progress on topics that matter within and beyond the academy.

Frédérique de Vignemont is a CNRS research director at the Jean Nicod Institute, Paris, France. Her research is at the intersection of philosophy of mind and cognitive science. Her major current works focus on the first-person, body schema, agency, empathy, and more recently on pain. She has just published *Mind the Body* (Oxford University Press, 2018), a philosophical exploration

of bodily self-awareness, as well as *The Subject's Matter* (MIT Press, 2017), an interdisciplinary collected volume on the body and the self. She is one of the executive editors of the *Review of Philosophy and Psychology*.

Dan-Mikael Ellingsen is a research fellow at the Martinos Center for Biomedical Imaging at Mass. General Hospital and Harvard Medical School, USA. His research focuses on the neural and psychological mechanisms of expectations, motivation, and clinician-patient communication in the therapeutic encounter, and how this shapes subjective experiences such as pain and positive affect. To investigate these questions, Dan-Mikael uses techniques such as functional and structural magnetic resonance imaging (MRI), measures of the autonomic nervous system, pharmacological stimulation, and self-report.

Matthew Fulkerson is Associate Professor of Philosophy at the University of California, San Diego, USA. He is the author of *The First Sense: A Philosophical Study of Human Touch* (MIT Press, 2014).

Valerie Gray Hardcastle is Professor of Philosophy, Psychology, and Psychiatry & Behavioral Neuroscience and Scholar-in-Residence at the Weaver Institute for Law and Psychiatry at the University of Cincinnati, USA. She studies the nature and structure of interdisciplinary theories in the cognitive sciences and has focused primarily on developing a philosophical framework for understanding conscious phenomena responsive to neuroscientific, psychiatric, and psychological data. Currently, she is investigating the neuroscience of violence and its implications for both our understanding of human nature and the criminal justice system. She received a bachelor's degree with a double major in philosophy and political science from the University of California, Berkeley, a master's degree in philosophy from the University of Houston, and an interdisciplinary PhD in cognitive science and philosophy from the University of California, San Diego, USA.

Colin Klein is Associate Professor in the School of Philosophy at the Australian National University. From 2014 to 2018 he was also an Australian Research Council Future Fellow. His work focuses on philosophy of science and philosophy of mind, especially where they intersect in philosophy of psychology. His book *What the Body Commands: The Imperative Theory of Pain* (MIT Press, 2015) won the 2016 David Harold Tribe Award in Philosophy.

Morten L. Kringelbach is Associate Professor in the Department of Psychiatry, University of Oxford, UK, and Professor of Neuroscience at University of Aarhus, Denmark. His prize-winning research uses whole-brain computational models of neuroimaging data measuring the hedonia (pleasure) of, for example, infants, taste, and music, in order to discover how to increase eudaimonia (the life well-lived). He has published 14 books and over 300 scientific papers, chapters, and other articles, and his research features regularly in newspapers, magazines, on radio, and on television. He is a fellow of The Queen's College, Oxford, and of the Association for Psychological Science, and he is on the

advisory board of Scientific American and is a board member of the world's first Empathy Museum.

Siri Leknes is Associate Professor of Social and Affective Neuroscience at the Department of Psychology, University of Oslo, Norway. Trained as a cognitive neuroscientist, Dr. Leknes' research centres around subjective feelings of good and bad, pleasure and pain. Specifically, Dr. Leknes asks how these feelings are produced by the peripheral and central nervous system in healthy humans, as well as in clinical conditions such as addiction and chronic pain.

Manolo Martínez is a Ramón y Cajal researcher at the philosophy department of the University of Barcelona Catalonia, Spain. His work focuses on issues in the intersection of the philosophy of mind and the philosophy of biology.

Jesse Prinz is Distinguished Professor of Philosophy at the Graduate Center of the City University of New York, and a Research Professor at the University of North Carolina, Chapel Hill, USA. Prinz has research interests in cognitive science, philosophy of psychology, philosophy of language, moral psychology, and aesthetics. He has written five books: *Furnishing the Mind: Concepts and Their Perceptual Basis* (MIT Press, 2002), *Gut Reactions: A Perceptual Theory of Emotion* (Oxford University Press, 2004), and *The Emotional Construction of Morals* (Oxford University Press, 2007), *Beyond Human Nature* (Penguin; Norton, 2012), and *The Conscious Brain* (Oxford University Press, 2015). He has published numerous articles on concepts, emotions, morals, consciousness, and other topics.

Adam Shriver is Research Fellow at the Oxford Uehiro Centre for Practical Ethics and the Wellcome Centre for Ethics and Humanities, Oxford University, UK. His research examines the intersection of ethics and cognitive science with a focus on human well-being and animal welfare. In particular, Adam's research has examined the significance of the dissociation between the affective and sensory components of pain for applied ethics and philosophical accounts of well-being. To this end, Adam has written about the relationship between pain and pleasure, the legal and ethical questions that arise from the search for a neural signature of pain in humans, and the capacity for suffering across different species.

Introduction

David Bain, Michael Brady, and Jennifer Corns

Over recent decades, pain has received increasing attention as – with ever greater sophistication and rigour – theorists have tried to answer the deep and difficult questions it poses. What is pain’s nature? What is its point? In what sense is it bad? The chapters collected in this volume are a contribution to that effort.

Understanding pain requires addressing two of its most obvious but least understood aspects: its unpleasantness and motivational force (how, that is, it drives us to behave in certain ways, for example, withdrawing your hand from hot water). Pain’s unpleasantness and motivational force seem closely intertwined. But what is the nature of each, and what exactly is their relationship? These questions are the focus of Part I.

A second route to understanding pain is through its connection with emotion, on which Part II focuses. For one thing, there appears to be overlap in the neural activity subserving pain and emotions. For another, just as pain is typically unpleasant, so are certain emotions. Indeed, pain’s unpleasantness may *involve* certain emotions. What, then, is the relationship between pain and emotional suffering? In what ways are they the same? In what ways different?

Finally, the atypical can sometimes illuminate the typical; hence Part III addresses deviant pain. Chronic pains, for instance, appear not to be associated with any physical damage, except perhaps to the nervous system. Pains undergone by pain asymbolics appear not to be found unpleasant. And the masochist’s pains seem to be wanted, even pleasant. Thinking about such cases offers insights into pain’s nature.

Below we briefly sketch the volume’s three overarching themes, and the contributions that follow.

Part I: Pain, unpleasantness, and motivation

Typical pains are unpleasant and motivational. The water becomes hot, you begin to feel pain, its unpleasantness increases, and you lift your hand out. Questions about pain’s unpleasantness and motivational force are beginning to receive the attention they deserve. In virtue of what are pains unpleasant? In virtue of what do they motivate us? What mechanisms underlie their unpleasantness and

motivationality? What influences these features? These questions are the focus of the chapters in Part I.

The first two chapters are moves in a philosophical debate about what makes pains unpleasant. Taking pain to be motivational because unpleasant, many explain pain's unpleasantness in terms of things that will also explain its motivational. The traditional view, for instance, appeals to experience-directed desires: your pain's being unpleasant is simply a matter of your *disliking* it in the sense of *wanting* it to stop for its own sake. But others have recently proposed alternatives. Some say pains are motivational (and unpleasant) in virtue of not desires but the possession of special intentional contents: imperative contents (according to imperativists) by dint of which your pains *tell you what to do*, or evaluative contents (according to evaluativists) by dint of which your pains evaluate bodily conditions *as bad* for you. Others reject such intentional accounts and claim, for example, that your pain's unpleasantness and motivationality turn not on its content as such, but on the kind of processing its content undergoes.

In the first chapter, 'Imperativism and pain intensity', Colin Klein and Manolo Martínez defend imperativism. A pain, they think, motivates in virtue of a part of its phenomenal character that consists in possession of imperative (rather than indicative) content. Your pains, in short, motivate because they are commands from your body: for you to *see to it that a certain bodily state not exist* (on Martínez's version) or for you to *protect a certain part of yours* (on Klein's).

This view, some worry, cannot capture pain's intensity, since the intensity of a command is not a matter of content. Bruiser might shout and swear when telling you to stand up, whereas Petal might speak softly and say 'please', but that difference is not an intentional difference.

Klein and Martínez's response is twofold. First, they distinguish command intensity from such phenomena as the politeness with which a command is expressed and the relative urgency of one command vis-à-vis others. Second, they use a possible-worlds model of imperative content to explain command intensity as an intentional phenomenon. The basic model says an imperative's content is the set of worlds at which it is satisfied. To incorporate intensity, Klein and Martínez argue, the content must include a function that *ranks* worlds. One imperative content, A, might be more intense than another, B, because A and B each rank worlds at which A is satisfied at B's expense over worlds at which B is satisfied at A's. With this idea in place, Klein and Martínez go on to propose more complex functions for more difficult cases; they also consider the bearing of their semantics on whether one pain might be not merely *more* intense but *twice as* intense as another; and they gesture at what it might be in virtue of that an experience possesses such a content – at, that is, a psychosemantics.

In Chapter 2, 'Pain and theories of sensory affect', Murat Aydede and Matthew Fulkerson reject both imperativism and evaluativism in favour of the idea that a pain's unpleasantness and motivationality consist in the kind of processing its content undergoes.

Among their objections to imperativism and evaluativism, a key complaint is that neither view explains pain's intrinsic badness. Consider an allodyniac

who is being caused agony by innocuous caressing of her forearm. Evaluativists say her experience represents to her (falsely) that her forearm is in a bad state. But Aydede and Fulkerson argue that this fails to explain why she has reason to stop that experience, hence to stop the caressing. Why should representations of badness themselves be bad? Imperativism, they argue, struggles with similar problems. For why should receiving a command from the body itself be bad? And even setting aside these normative questions, can imperativism explain pain's motivational? Commands, Aydede and Fulkerson suggest, are not inherently motivational. When the police command a crowd to disperse, the crowd might be unmoved. Pain, they argue, is better compared to water cannon.

Instead, Aydede and Fulkerson propose a psychofunctionalist account: a pain is unpleasant in virtue of the 'm-processing' undergone by the sensory information it carries (about bodily states and events). M-processing is something that happens subpersonally, they explain, hence its precise contours should be delineated by science; but Aydede and Fulkerson anticipate that m-processing will be an inherently motivational mode of processing, specified in terms of the processed information's effects on the organism's preferences, motor systems, and learning. Hence their overall picture is that an unpleasant pain has a belief-like component, which carries sensory information about a bodily condition, and a desire-like component (consisting in that information being m-processed), which is not identical to but can be *modelled* on a personal-level desire for that represented bodily condition to cease. And it is in these terms that they explain unpleasant pain's intrinsic badness. Its badness, they argue, consists in the frustration of the desire-like component 'as registered by' the belief-like component.

The final chapter of this section presents 'A neuroscience perspective on pleasure and pain'. In this, Dan-Mikael Ellingsen, Morten L. Kringelbach, and Siri Leknes address unpleasure and pleasure – or affect – *in general*. As neuroscientists, their approach is rather different from the preceding chapters'; but there are threads of continuity, not least their interest in the relationship between affect and motivation.

The relationship between affect and motivation is, they suggest, complex. Distinguishing between 'liking' (manifest in feelings of pleasure and such facial expressions as lick-lipping) and 'wanting' (manifest in approach behaviour, for instance) the authors report cases where 'wanting' seems to be increased independently of 'liking'. In one, injections of dopamine into rats' brains made the rats work harder for a sugary treat yet not lick their lips more once they got it; in another, human subjects pressed a lever obsessively to cause brain stimulations it wasn't clear they were enjoying.

There are also, the authors argue, cases showing the looseness of the relationship between experiences' sensory and affective components. Some are ordinary, as when chocolate is found pleasurable at first but disgusting after over-eating. Others are extraordinary, for instance pain asymbolics failing to find even *pain* unpleasant. There are also cases in which experiences seem *both* pleasant and unpleasant, the authors argue, as when winners of sub-optimal prizes say they feel

both good and bad, putting pressure on the idea of a single continuum between pleasure and displeasure.

Turning to the diverse and flexible *determinants* of affect, the authors argue that these are complex, instructive, and (despite the ‘liking’/‘wanting’ distinction) tied to motivation. Homeostatic needs can seem to be key determinants, as when sodium-depleted rats display strong ‘liking’ reactions to intensely salty water. But context is also important. On one theory, your pain’s unpleasantness will be reduced if you unconsciously deem what’s causing it less important than, say, an impending threat. Expectations, associations, and the ‘meanings’ attached to events can all, relatedly, play a role – as demonstrated, the authors argue, by a range of cases, some involving placebos (see also Corns’ chapter), others involving subjects finding gentle touch more pleasurable when the toucher is attractive. Such flexibility between stimuli and affect, the authors conclude, suggests that affect’s role is less to provide us with information than to motivate us, on the basis of unconscious calculations, towards the best actions.

Part II: Pain and emotion

The relation between pain and emotion is a fruitful area for new perspectives on affective experience. It has been long thought that physical pain has an emotional element: consider for instance the famous IASP definition of pain, first suggested in 1964, that it is ‘[a]n unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage’.¹ Moreover, as Jesse Prinz points out in his contribution, philosophers as wide-ranging as Aristotle, Spinoza, and Hume all hold that emotions are *forms* of pain or pleasure. This chimes with common-sense thinking: it is widely held (albeit sometimes metaphorically) that some forms of emotional suffering are themselves painful – we talk of the pain of a broken heart and of disappointment, of being pained by grief and loss. Despite these close connections, philosophical research has seldom addressed both forms of negative affective experience together, preferring instead to focus on one form and address the other only in passing. This is both surprising and unfortunate: surprising, given that links between the two seem well-known and rather obvious; unfortunate, given that research and progress in our thinking about one kind of experience might profitably inform our thinking about the other. A focus on pain or emotion to the exclusion of the other thus threatens to be detrimental to philosophical progress in either area.

The chapters in this second section all deal, to a greater or lesser extent, with points of comparison and contrast between pain and emotion. Michael Brady’s chapter – ‘The rationality of emotional and physical suffering’ – focuses on the extent to which pain and negative emotion are reason-responsive. Brady wishes to defend a traditional view according to which forms of emotional suffering – like grief, disappointment, and shame – are reason-responsive, and thus assessable as rational or irrational, whilst forms of physical suffering – like pain and hunger – are not. He thus argues that although there are many ways in which physical and emotional suffering are alike, they differ considerably at the normative level.

Brady makes his case by arguing that there are a number of reasons to think that forms of emotional suffering are reason-responsive; he cites introspective, meta-ethical, and developmental evidence. The first kind of evidence suggests that we need to appeal to evaluative content in order to distinguish emotions from each other, but that such content is not necessary for introspection to distinguish different kinds of physical suffering. The second focuses on the idea that the values to which emotions respond are best understood on 'rational sentimentalist' lines, that is, in terms of features that make the emotion rationally appropriate. The objects of physical suffering do not admit of a plausible rational sentimentalist treatment, however. Finally, he argues that there are good developmental reasons for forms of emotional suffering to involve an evaluative stance; in particular, this is needed to accommodate the wide range of objects that trigger emotions and flexibility in behavioural response. Neither of these reasons, he argues, are applicable when we consider forms of physical suffering. He maintains, in light of this, that there are good reasons to think that emotional suffering is, and physical suffering is not, reason-responsive, because emotional suffering is, and physical suffering is not, imbued with evaluative content. Brady concludes, finally, that this puts pressure on evaluativist accounts of pain and other forms of physical suffering, since these hold that pain and physical suffering do have evaluative content.

In 'The placebo effect', Jennifer Corns looks at another point of comparison between physical and emotional suffering, namely the susceptibility of each to the placebo effect. Corns begins her chapter by pointing out an interesting point of dissimilarity between physical and emotional suffering: the former seems susceptible to placebo effects whilst the latter is not. Corns thinks this dissimilarity is troubling for attempts to characterize the placebo effect and proceeds to argue that there is no clear characterization available and none likely to succeed, whether for pain or emotion.

Corns makes her argument by considering six ways that non-placebo effects have been distinguished from the placebo effect in the existing literature. These are (i) active versus inactive, (ii) real versus fake, (iii) pharmacological versus psychosocial, (iv) specific versus non-specific, (v) treatment effects versus context effects, and (vi) legitimate healing versus illegitimate healing. In each case, Corns argues that there are considerable problems with drawing the purported distinction in this way. This would seem a troubling conclusion if it was important to have a successful characterization of the placebo effect. But Corns goes on to argue against the utility of identifying any class of effects as placebo effects. In particular, such an identification does not help in the areas where it is claimed to help, namely in (i) randomized control trials, (ii) inquiries into the supposed mechanisms of the effect itself, and (iii) treatment. Finally, Corns explains away the intuitions we have about there being a distinct class of placebo effects and argues that doing so can help to improve our treatment of all kinds of suffering. Far from being useful to our medical practices, therefore, the idea that there is such a distinct class can hamper effective treatment.

Jesse Prinz's chapter, 'What is the affective component of pain?', focuses on one of the central ways that pain and emotion are thought to be connected,

namely (as mentioned earlier) the idea that pain has both a sensory component and an affective component. Common sense and scientific theory support this 'componential' view of pain, with the sensory component purportedly carrying information about the body – e.g. that it has been damaged – and the affective component corresponding to the feeling of badness or unpleasantness that makes pain something to avoid. It is Prinz's aim to investigate the nature of the affective component. He begins by looking at accounts which identify the affective component with an emotion, first considering and dismissing the idea that there is a distinctive emotion characteristic of pain, and then arguing against the claim that one or more familiar emotions (such as anger and fear) constitute the emotional element. Prinz proposes, instead, that negative affect can be identified with *negative valence*. This is not an emotion, although it is a component of some emotions.

Prinz then reviews different theories of valence and proposes his own account, according to which valence is an 'inner marker' that (in the case of pain) tells us to act so that the sensory component ceases or stops. After defending this account by showing that it can accommodate a wide variety of empirical findings, Prinz then discusses a potential objection to his account of valence as an internal marker; since such markers are not on his view to be identified with feelings, then the affective component of pain that is constituted by negative valence is not itself an unpleasant feeling: it does not feel bad, in other words. Although this sounds paradoxical, Prinz concludes by explaining reasons for doubting that pains necessarily involve an unpleasant feeling in addition to their sensory component. This means that we should reject a 'folk platitude' about pain: namely, that it is necessarily unpleasant; but, Prinz thinks, doing so can lead us to a clearer philosophical view as to why pain is after all bad.

Part III: Deviant pain

Paradigmatic pains hurt, are felt as being located in one's body, are associated with damage or injury, are motivational, conscious, and more besides. But many of the pains that we experience are not paradigmatic. A good theory of pain should be able to account not only for paradigmatic pains, but also atypical, deviant pains. Though deviant pains make an appearance in the first two parts of the volume, each of the chapters in Part III provides a targeted, empirically rich discussion of a type of deviant pain and its importance for philosophical theories of pain.

In 'The unpleasantness of pain for humans and other animals', Adam Shriver takes up the most discussed case of deviant pain within philosophy: pains that subjects report not being bothered by. The standard interpretation of these deviant pain cases is that the affective component of pain is absent, while the dissociable sensory component remains.

After giving an overview of the evidence for this type of deviant pain, Shriver explains its philosophical relevance. In the philosophy of mind, for instance, change in the affective component without a change in the sensory component

may be thought to raise problems for representational theories of pain which require phenomenal qualities to supervene on intentional content. In ethics, for instance, our understanding of the relationship between the two components arguably has implications for what we think is ultimately valuable or disvaluable. Dissociations raise further questions in applied ethics about non-humans and developing humans who may only experience pain's sensory component.

Shriver claims that though philosophical questions like these cannot be settled by scientific inquiry, neither can they be settled without it. He first turns to research aimed at identifying affective pain processing in non-humans, but points out how such research is surprisingly limited due to difficulties separating behavioural reflexes from affect. Accordingly, he argues for a three-pronged approach combining what we can learn about the relevant neural mechanisms in humans with both behavioural evidence and drug reactivity from humans and non-humans. Shriver considers this evidence and ultimately concludes that it is currently inconclusive.

Since answering the philosophical questions raised by sensory/affective dissociation is limited by these gaps in our scientific understanding, Shriver spends the remainder of the article suggesting ways in which philosophers and scientists have and might further work together to fill them. In particular, he focuses on existing research into correlations between unpleasantness and both learning and motivation. Throughout this discussion, he raises further philosophical questions and experiments that he thinks could advance our understanding.

In the next chapter, 'When is a pain not a pain? The challenge of disorders of consciousness', Valerie Gray Hardcastle focuses on the pains of patients suffering from disorders of consciousness. The pains experienced by these non-conscious, or minimally conscious, patients may be understood as another type of atypical, deviant pain, consideration of which Hardcastle thinks helps to reveal what is and is not morally significant about pain and appropriately targeted for treatment.

With some reservations about the term, Hardcastle provides an overview of *disorders of consciousness*. Such conditions range from the clearly conscious people suffering from locked-in syndrome to the clearly non-conscious coma patients. Hardcastle focuses mostly on the middle of this spectrum and describes the tests – both behavioural and neural – used to determine how conscious such patients might be. Hardcastle is sceptical about what this evidence reveals and, more importantly, thinks the results are irrelevant for whether these patients are in pain.

It is, Hardcastle argues, pain processing that matters – both morally and for treatment; in particular, it does not matter whether the person is experiencing conscious pain or whether the person is conscious. Whether pain is being processed can be ascertained independently of whether the patient is conscious. Hardcastle argues that consciousness is not necessary for moral worth; in particular, it is not necessary for the two classic contenders for moral standing: being rational and having interests. Moreover, against a more traditional model invoking dedicated pain pathways, Hardcastle argues that pain processing is best understood as being carried out by massively parallel processing across a complex

neural network. She then compares the activity in this network between healthy subjects and patients suffering from various disorders of consciousness. Her aim is not to adjudicate this evidence, but instead to convince us that it is the processing of information across this network that matters, and not whether any of this processing is conscious. The final section of Hardcastle's chapter focuses on the negative effects of pain processing as evidenced in both premature infants who have undergone multiple painful procedures and chronic pain patients. She concludes that not treating pain is detrimental to the brain and central nervous system – and that this is as true and as morally relevant for the deviant pains undergone by those suffering from disorders of consciousness as it is for paradigmatic pains.

In the final chapter, 'The first-person in pain', Frédérique de Vignemont focuses on the relationship between pain and bodily ownership. She argues that for a *pain* to be felt as yours, it is not necessary that the felt pain be localized in a *body* felt as yours. Pains felt in such an 'alien' body – a body that you feel, but not as yours – are clearly another type of deviant pain.

De Vignemont presents some of the reasons that it has seemed that pain and bodily ownership are instead conceptually inseparable, i.e. reasons that have been offered for the claims both that I necessarily feel my body as my own whenever I have bodily sensations (including pain), and that I necessarily feel a body part as my own when I localize pain within it. Consideration of these reasons gives rise to two key hypotheses: that being able to feel pain and feeling that something is a part of my body are jointly sufficient for localizing pain in that something ('the sufficiency claim') and that it is necessary that something be felt as part of my body in order for me to feel pain within it ('the necessity claim'). De Vignemont tests these hypotheses against three types of cases: the Rubber Hand Illusion, ownership delusions, and disownership syndromes. She argues that while the sufficiency claim is consistent with these cases, the necessity claim is not and should therefore be rejected.

While pain and bodily ownership thus dissociate, de Vignemont argues that *threat* and bodily ownership do not: to experience a bodily threat, I must experience the threatened body as my own. She ultimately explains this difference by invoking the different spatial organization of the two types of experiences: a threat is always more or less external to the body, but pains are not. Threats, unlike pain, are therefore essentially tied to sensing the boundaries of one's body. De Vignemont concludes by noting that even if there are rare, deviant cases in which we feel pains in alien bodies that we do not feel as our own, we nonetheless always feel the pain as our own.

This volume arose from what came to be known as the Pain Project. Operating under the aegis of the Pain and the Nature of Minds programme at the University of Notre Dame, the Pain Project focused on relations among pain, perception, and emotion, and on pain in non-human animals. Its core team – Principal Investigators David Bain and Michael Brady, and their postdoctoral fellow, now lecturer, Jennifer Corns – are philosophers at the University of Glasgow. The broader team, based in Glasgow, Paris, and Oslo, comprised philosophers of mind

and cognitive science, ethicists, neuroscientists, and veterinary scientists. Over 18 months, from January 2012 until June 2013, the Project ran four workshops and a conference, at which most of the contributions to this volume originated. The present volume exhibits the progress being made in understanding pain. But it is also clear that further progress is needed. There are many questions yet to be answered. We hope the chapters collected here stimulate further work into the nature of pain.

Note

- 1 'Part III: Pain Terms, A Current List with Definitions and Notes on Usage', *Classification of Chronic Pain*, Second Edition, IASP Task Force on Taxonomy, edited by H. Merskey and N. Bogduk, IASP Press, Seattle (1994), pp. 209–214.

Part I

Pain, unpleasantness, and motivation

1 Imperativism and pain intensity

Colin Klein and Manolo Martínez

1 Introduction

Pains vary in intensity. A good philosophical theory of pain should say something about those variations. For one, variations in the intensity of pain are, from a practical standpoint, almost as important as the presence or absence of pain itself. An ibuprofen may dull but not eliminate the pain of your sore muscles. That will have consequences for how you feel, what you do, and what your pain is like.

For another, if you *don't* say something about degrees of pain, your philosophical opponents will tend to assume that you *can't*. We have both recently defended versions of an *imperative* theory of pain.¹ On such a view, pains have imperative contents that express commands to do or avoid certain actions. For example, in Martínez's version, the content of pain experiences is analogous to:

See to it that this bodily state does not exist!

Similarly, in Klein's more recent formulation (2015), a pain is an imperative along the lines of:

Protect this body part (in this manner).

On both accounts, a pain is an imperative that would be satisfied if the sufferer of pain ceased to be in a particular bodily state. That bodily state is often one that is threatening to life or health. On the whole, then, pains command us to do things that will keep our bodies intact and well-functioning. Not all pains need do so—the system might misfire in a variety of ways—but the *function* of the pain system is to keep us healthy.

Pains work this way because of the nature of imperatives. Imperative content prescribes some action to be taken, or some goal state to be attained. Pains are not the only sensation with imperative content. The difference between pain and many other sensations can be explained in terms of the differences in the imperatives that constitute them (Hall 2008). Itches command *scratching* regions rather than protecting them. Hunger commands *eating* rather than an action directed at a particular body part. Finally, imperative modalities differ from representational modalities like vision and touch: the latter *represent features* of the world and

have truth-apt contents, while the former *command an action* and so do not have truth-apt contents.²

In defending imperativism, neither of us has said much about pain intensity. Some have suggested that this is because there is nothing to be said: imperativism, for principled reasons, cannot give a story about intensity. That would appear to be a point in favor of representationalism about pain. After all, many facts about the world vary in magnitude. If pains represent something, then it is easy to see how that something might vary in magnitude, and thereby how pains themselves might have different intensities. The imperativist—who does not think that pains are in the business of tracking properties in the world—has no corresponding facts to appeal to. At least, that's how the objection goes.

Our opponents are too pessimistic. We can give a perfectly satisfying account of imperative intensity. We'll present the case in two steps. First, we must establish the general fact that imperatives can vary in intensity. We'll do that in Section 2. Along the way, we will respond to several critics of the imperative theory. Second, we must give a model for variations in intensity, in which the variations are part of the *content* of the imperative. Imperativism is most exciting as an intentionalist theory, one in which phenomenal properties supervene on intentional contents. Failure to reduce variations in intensity to variations in content would weaken the appeal. We give such a model in Section 3 and conclude with a few brief reflections on the naturalization of imperative intensity.

2 Imperatives come in degrees

2.1 Imperative intensity

Commands come in degrees. All things being equal, shouting:

Pass the salt now!

conveys a more urgent command than does the calm:

Please pass the salt at your earliest convenience!

We will call these variations in the *intensity* of the imperative. Variations in imperative intensity have practical consequences. All things being equal, a more intense imperative should be more likely to make you change your plans, to perform the commanded action sooner, to weight the commanded action higher when deliberating among mutually exclusive courses of action, and so on.

More formally, we will assume that the content of an imperative can be partially identified with a set of *satisfaction conditions*, here modeled as a set of possible worlds W_{sc} .³ “Pass me the salt!” is satisfied just in case you pass me the salt. So the content of the imperative is, at a first pass, the set of worlds in which you pass me the salt. A command from a source that you take to be legitimate gives you a reason to act so as to bring about the satisfaction conditions. Such commands will thus have practical consequences for your actions.

We will treat the intensity of a command as a further weighting on the set of satisfaction worlds. At its most general, then, the proposal is this: the content of an imperative includes both a set of satisfaction worlds and a degree of intensity. Intensity simply functions as a weighting across possible worlds, dividing them into better and worse, in ways to be elaborated shortly.

Intensities might be relatively vague, as in the case of the salt commands above. Or the degree may be more precisely specified, as it is in certain more formal contexts: triage classifications in emergency rooms, the rubber stamps used by enthusiastic managers to prioritize their edicts, or the priority codes on the Autovon telephone network.

The Autovon priority system is an especially nice example of what we have in mind. The Autovon network was developed during the Cold War to provide communication during a nuclear attack. An Autovon keypad had an additional column of four keys that allowed the user to specify the precedence level of the call. The dialed number and the precedence level thus ordered the network to route the call to a certain location, with a certain precedence. (Note that the priorities relate to the *routing of the calls*, not to the contents of the calls themselves. The White House could use Autovon to order a pizza during a nuclear war. While the call itself would trump all others, delivery wouldn't be high on anyone's list of priorities.) Higher precedence calls would, if necessary, kick lower precedence calls off of the trunk to ensure that a call went through. The highest level, which would guarantee that a call would go through if any could, was restricted to the White House.⁴ We will argue that imperative intensity can be generally modeled by something like the Autovon's ranking system. Before we get there, however, we want to say a few words about what imperative intensity is *not*.

2.2 What intensity isn't

That commands *can* vary in intensity is, we take it, obvious enough. Further, the extension to pains seems straightforward enough: pains are imperatives; like other imperatives, they can vary in intensity, and that variation has exactly the same sorts of practical consequences. A more intense pain in your foot will, all things being equal, cause you to forgo more activities in order to tend to your foot, to tend to your foot more urgently, and so forth.

There is still work left to do to ensure that intensity can be usefully modeled. Before we do it, though, it's worth distinguishing intensity from a few things that look a lot like intensity but aren't. Failure to do so has, we think, led to some of the objections to the imperative account.

First, the intensity of a command ought to be distinguished from the *illocutionary force* of an *utterance* of an imperative. Imperatives can be polite or nasty; they can be phrased as requests or pleas or straightforward commands. None of these are variations in intensity *per se*.

Failure to distinguish the two may lead to the impression that variations in intensity are not part of the content of imperatives. In a recent paper, Cutter and Tye have objected that an appeal to intensity amounts to abandoning

intentionalism. Regarding the proposal that imperatives vary in intensity, they write that:

On such a proposal, the difference between [two pains] is analogous to the difference between the following two imperative sentences:

- (Please) stop that bodily disturbance.
- Stop that bodily disturbance!!!

(Cutter & Tye 2011, p. 104)

They then object that such an account is inconsistent with intentionalism because:

[N]ow the phenomenal character of an experience does not supervene on its content alone; rather it supervenes on its content together with its degree of urgency.

(Cutter & Tye 2011, p. 104)

We agree with Cutter and Tye that the two sentences above may not differ in content. The variations in the two sentences above are variations in *politeness*, however, not the intensity of the imperative.⁵ More generally, two imperatives may differ in illocutionary force without differing in content. The very same command may be ordered, demanded, requested, or politely suggested. Why? Because commands are individuated by their content—on our account, the satisfaction-conditions of the command, plus a weighting over worlds—while differences in illocutionary force depend on the social circumstances under which imperatives are uttered. So the very same command, expressed in varying external circumstances, might do different things without changing in content.

There is the potential for confusion, we suggest, because sometimes facts about intensity can be expressed by uttering imperatives with the right illocutionary force—rude commands tend to express high-intensity imperatives, for example, and polite ones low intensity. But note that that's entirely compatible with intensity being part of the *content* of the command. Compare: I might express the relative temperature of a 110-degree day by using a variety of choice expletives. The vigorousness of my swearing might accurately convey the degree to which some external magnitude varies. Yet my claim about the temperature still depends, in an important sense, on what I have said. So too with imperatives.

Given the potential ambiguity, though, we will separate *commands* (a certain type of content) from *orders* (a type of speech act that can convey a command). Note more generally that commands might need to be distinguished from related contents like *permissives*, which permit but do not obligate action (Hamblin 1987).

Second, intensity ought to be distinguished from the *temporal priority* or *urgency* of the command. Some imperatives ought to be satisfied sooner than others. This temporal ordering may depend in part upon the intrinsic content of an imperative. However, priority itself seems like an *extrinsic* property of imperatives. That urgent bit of dusting becomes less urgent when the house is on fire; the urgent

ache in your toe suddenly becomes less so when the bear appears from behind the tree. Of course, priority depends on intensity: all things being equal, a more intense imperative should also be satisfied sooner. Our account will preserve this feature. But it also depends on what else is going on, and to which other imperatives you're subject. So while imperative priority is tightly related to intensity, it is not the same thing.⁶ Hence, while urgency is extrinsic, intensity need not be.

Third and finally, imperative intensities are features internal to imperatives. They need not track any magnitude in the world. That should not be surprising: the imperativist account, remember, says that pains are primarily spurs to action rather than states that track the world.

Adam Pautz uses this fact to object to the imperative theory. He writes:

While negative imperatives admit of degree, it is hard to see how their degrees might match up with degrees of painfulness. What in the imperative contents of [an agent's] two consecutive pains ... might determine that the second pain was roughly twice greater than the first?

(Pautz 2010, p. 364, fn. 36)⁷

We think that this is to misunderstand the imperativist position. For all we've said, there's a simple answer to this question: what determines whether one pain is twice greater than another is simply that it is twice as intense. That answer will need some refinement, but in general—the fact that imperatives come in degrees opens the possibility of comparing those degrees.

Of course, pain intensities differ from the intensities of sensations like light and sound. The latter tracks changes in the magnitude of something in the world. The former doesn't. Insofar as judgments about pain intensity are about anything, they are about *pains*—and their effects on our motivational state—rather than about the world.⁸ But this seems like an attractive feature of our account, not a problem. The intensity of pain tracks intensity of (say) injury—but only very roughly. Pain from the very same injury may ebb or fall in intensity through the day as we rest or take painkillers, or simply by its own obscure logic. We aren't inclined to think that the *injury* has changed, only that the *pain* has. More generally, it's always an open question whether changes in the intensity of pain correspond to some change in our injuries. It is not always clear to us whether our pains correspond to any injury at all. Indeed, there is an extensive empirical literature, in the form of the gate control theory of pain and its successors, devoted to elucidating the mechanisms of this dissociation.

So there is no obvious problem here. There is, however, a subtler problem lurking in the vicinity. Subjects will consistently *judge* that two pain intensities are ratio multiples of each other in laboratory settings (Price et al. 1983). This is often taken as showing that pain intensity itself is a proper psychophysical magnitude: that is, that it has a theoretically motivated zero point (Nunnally 1967) and that it makes sense to be realists about subjects' judgments because they pick out actual ratios of a real quantity. If so, then pain intensity is comparable to (say) felt intensity of loudness or brightness.

However, the question of whether pain intensity is really a true magnitude, and so *properly* measured by a ratio scale, is a contentious question. The alternative would be that pain intensities might be merely *orderable* rather than true magnitudes. That is, it would make sense to talk about pains being stronger or less strong than one another, but not for one pain to be twice as bad as another.⁹ The mere fact that subjects can respond to ratio questions and that their responses fit a power curve does not itself show that the scale is a ratio scale. For as Hall notes, this is a task “which *presupposes* that subjects can judge ratios” (1981, p. 103). The fact that the resulting judgments fit a power function (especially across the limited range available to ethically acceptable laboratory experiments of pain) may thus show “nothing more than the flexibility of the power function as is evident in its capacity to fit any of the monotonically increasing curves common-sense would expect to describe the relationship between numerical judgments and stimulus intensity” (p. 104). Ordinal pain intensity would not necessarily prevent people from making judgments about ratios of pains *when instructed to do so*. That might even come out as consistent, especially across limited ranges of stimuli (such as are ethically inducible in a laboratory). Nevertheless, the ultimate question is not which scale *works* but which scale is appropriate for measuring pain.

We think it wise to sidestep this debate. Our goal in the next section, when we model pain intensities, will be only to give a theory that is consistent with subjects’ *judgments* of pain intensities, rather than one which assumes something more controversial about the structure of what underlies those intensities themselves. That should make the resulting theory compatible with both magnitude or ordinal theories of pain intensity. It’s worth noting at the outset, however, that magnitude theories get our proposed structure on the cheap. Magnitudes can be represented by simple cardinal numbers on a scale with a true zero. Those should always permit the sort of inter-comparisons that we will propose. If we were sure that pain intensity were a magnitude, it would be tempting just to represent it by a simple number and move on. Since this isn’t obvious, it’s worth building an account that works even on ordinal theories. Further, even if *pain* intensity happens to have a simple structure, *imperative* intensity more generally probably needs an ordinal treatment, so the account should be of broader interest.

3 Intensity and content

3.1 Preliminaries

Imperatives vary in intensity; so do pains. A natural conclusion: variations in intensity of pains are due to variations in the intensity or degree of the imperatives that constitute them.

That is not quite enough, however. As we’ve emphasized, imperativism is a species of intentionalism. Intentionalism claims that phenomenal content supervenes on the content of experiences. We must therefore show that intensity is part of the content of imperatives.

That is not so obvious. One might hold, for example, that imperative intensity is an extrinsic property (like urgency). We have said this is distinct, but that's far from obvious. If it wasn't, that wouldn't necessarily be fatal to imperativism: one could still claim that pains are imperatives but that their intensity depends on their merely functional (rather than content-bearing) relations to other mental states. That would be a sort of quasi-intentionalism, which might in turn still be amenable to a naturalistic treatment.¹⁰ Nevertheless, we take it that one of the attractions of imperativism is its ability to link content and consciousness in what have been traditionally hard cases for the intentionalist, and so it's worth trying to preserve that feature.

The imperativist must also avoid the appearance of an *ad hoc* solution. So, for example, one might treat the content of a pain as simply the ordered pair:

$$(W_{sc}, d)$$

where W_{sc} is the set of worlds in which the imperative is satisfied, and d is the intensity. That would solve the surface problem neatly. But without saying more, it would be under-motivated. Ideally, the imperativist ought to get the intensity of pain to fall out of a more general theory of imperatives and their structure.

In the remainder of this section, we sketch a model of imperatives in which intensity falls out as a natural, intrinsic part of content. We will then return to pain, using our framework to explain some puzzling features of pains, and reflecting a bit on the naturalization of intensity.

3.2 Imperatives and ranking

Remember that we initially identified the content of an imperative with the set of possible worlds in which the imperative would be satisfied. This can't be sufficient, however. The command:

(O) Raise money for Oxfam!

admits of two readings: one in which you are commanded to raise *any* amount of money, another in which the more money you raise, the better. Both imperatives would be satisfied in the same worlds: namely, ones where I raise any amount of money for Oxfam. But the two clearly express different commands. This is because the second *ranks* satisfaction worlds: the more money I raise, the more preferable the world. Similarly, "Clean your room!" expresses a different command from: "Clean your room; and the sooner, the better!" Both are satisfied in the same worlds. The latter contains an additional exhortation, though, which ought to get you moving faster.

This suggests that the content of an imperative is larger than just the satisfaction worlds: it is instead a set of satisfaction worlds plus a *ranking function* $\geq$ defined over the set W of all possible worlds. The ranking function fills the

gap identified above by introducing a mechanism whereby imperatives satisfied in the same worlds might be compared. For any two possible worlds, w_i and w_j , $w_i \geq w_j$ means that w_i is ranked at least as high as w_j . Two worlds are equally ranked if they are ranked at least as high as each other, and a world w_i is ranked strictly better than w_j just in case w_i is ranked at least as high as w_j and they are not equally ranked.

We can thus model the content of an imperative i as the ordered pair:

$$(W_{sc}, \geq)$$

For most simple imperatives, satisfaction is better than non-satisfaction. So each world in W_{sc} is ranked strictly better than each world not in W_{sc} , each pair of worlds in W_{sc} is equally ranked, and *mutatis mutandis* for worlds not in W_{sc} . So $\geq$ partitions W into exactly two equivalence classes: the satisfaction worlds and the rest, with worlds in the former preferable to the latter. On the first reading of (O), $\geq$ will simply define a partition with two equivalence classes: one that includes all and only those worlds in which the addressee of the imperative raises any amount of money for Oxfam, and another that includes every other possible world, and where any world in the former equivalence class is strictly preferable to any world in the latter.

More complex imperatives partition W into more equivalence classes. On the second reading of (O), $\geq$ will provide a more complex ranking—the worlds in W_{sc} are better than the worlds outside, but the worlds within will also be further partitioned into smaller equivalence classes, each composed by all and only those worlds in which the addressee of the imperative raises a certain amount of money.

This captures the subtle distinction between the two *meanings* of (O). That distinction is grounded in the differences in the $\geq$ of the two readings, not in the satisfaction-conditions. That suggests that $\geq$ is part of the content of imperatives, and necessary to make certain distinctions between them.

Given this expanded content, two observations are the key to a theory of imperative intensity.

First, note that the $\geq$ for an imperative i ranks over all possible worlds. This means that it will also include worlds in which *other* imperatives are satisfied.

Second, note that $\geq$ is not restricted to ranking worlds in W_{sc} as higher than worlds not in W_{sc} . That is, it is possible to have an imperative that ranks the satisfaction of some *other* imperative as strictly better than its own satisfaction. That may seem odd. But note that it is precisely the case with the formalized structures of imperatives noted in Section 2.1. An Autovon call with status IMMEDIATE ought to be routed to its destination. Further, it ought to be routed preferably to calls with status ROUTINE. So worlds where it is routed and a ROUTINE call is not are ranked as strictly better by $\geq$ than worlds in which the reverse is the case. And finally, the call ought to be *dropped* in case of conflict with a call of status FLASH. So there are non-satisfaction worlds (in which the call fails in favor of a FLASH call) that are ranked by $\geq$ as strictly better than satisfaction worlds (i.e. where the call goes through at the expense of the FLASH call).

We propose that this sort of structure is actually constitutive of imperative intensity. Say that an imperative *i* is satisfied *at the expense of* some imperative *j* just in case either *i* is satisfied and *j* is not or *i* is satisfied temporally prior to *j* being satisfied. Say that an imperative *i* is *semi-preferable* to an imperative *j* relative to some (possibly hypothetical) distinct imperative *k* just in case either:

- 1 Both *i* and *j* rank all worlds in which *i* is satisfied at the expense of *j* as better than worlds in which *j* is satisfied at the expense of *i*.
or
- 2 *i* and *j* cannot be satisfied at the expense of one another, but *k* is such that
 - a *i* ranks all worlds in which it is satisfied at the expense of *k* as better than worlds in which *k* is satisfied at the expense of *i* and (2) *j* does not rank all worlds in which it is satisfied at the expense of *k* as better than worlds in which *k* is satisfied at the expense of *j*.
or
 - b *i* ranks some worlds in which it is satisfied at the expense of *k* as not worse than worlds in which *k* is satisfied at the expense of *i* and (2) *j* ranks all worlds in which it is satisfied at the expense of *k* as worse than worlds in which *k* is satisfied at the expense of *j*.

Given this, we can define imperative intensity as follows:

Intensity: An imperative *i* is more intense than an imperative *j* just in case there is some imperative relative to which *i* is semi-preferable to *j* and no imperative relative to which *j* is semi-preferable to *i*.

Let's unpack that a bit. Clause one covers the simple—and by far the most common—cases where *i* ranks its own satisfaction as better than *j*'s, and *j* agrees. How do we know which of two Autovon phone calls is more important? Well, a FLASH call *i* ranks its own connection as better than some ROUTINE call *j*, and *j* also ranks its completion as less important than the completion of *i*. So *i* is semi-preferable to *j*. As that structure is consistent across the different Autovon levels, *j* is never semi-preferable to *i*. Hence *i* is more intense than *j*.

The third imperative *k*, and the indirect structure it makes possible, are necessary for a tricky subset of cases (ones, however, that pains arguably exemplify). Consider the imperative: "Repent of your sins, and the sooner the better!" It seems possible that the pastor and the prophet might issue these commands in a way that varies in intensity—the prophet utters it as a matter of gravest importance, while the pastor is more understanding of the complexities of modern life. Yet the two imperatives have the same satisfaction conditions W_{sc} . Further, they always rank earlier satisfaction as *ceteris paribus* better than later satisfaction. So it's not obvious how to drive a wedge between the two. Note that this problem appears whenever we have two imperatives with the same W_{sc} : it is not possible to construct worlds in which one but not the other is satisfied, or where one is satisfied earlier than the other, and so it's hard to see how they might be ranked.

That's the point of the second clause. For we can still construct semi-preferability by reference to *other* things one might do. Consider some third imperative k —say, to move your car out of the fire lane. The prophet cares not for fire lanes; i_{prophet} ranks repentance higher than satisfying k . The pastor understands the importance of rendering unto Caesar and all that. So i_{pastor} might rank satisfying k as more important than immediate repentance. (Note that the ranking stipulated by k is irrelevant: it is only the ranking of various forms of i relative to the *satisfaction* of k that is at issue.) This is consistent with the stipulation that the pastor wants you to repent sooner rather than later, so long as that is read with an appropriate *ceteris paribus* clause. For it is true that, everything else being kept fixed, earlier repentance will be better than later. If so, then i_{prophet} is semi-preferable to i_{pastor} . Assuming that this structure holds generally, then the prophet's imperative is more intense than the pastor's, even though both are satisfied in exactly the same conditions. Hence for very similar or identically satisfied imperatives, their intensity can still be compared by triangulating against *other* possible things you might do.

3.3 Advantages of our approach

This approach to imperative intensity has several advantages. Most happily, imperative intensity ends up depending on content. Furthermore, it does so by depending on content in a principled way: the inclusion of $\geq$ is necessary to capture the differences between the readings of imperatives like (O). Second, we have done so using a purely ordinal measure, $\geq$, which leaves open the possibility that the extension to pains might treat pain intensities as merely ordered.

The theory on offer also explains why imperative intensity makes the most sense in limited domains—or, conversely, why many imperatives have *incommensurate* urgencies. If I utter (O) and your department chair orders you to repaint your office, there may not be a well-defined sense in which one of those commands is more intense than the other. That falls out nicely from our theory. The two imperatives are mutually selfish: each ranks its own satisfaction worlds higher than worlds in which it fails to be satisfied and the others are. So neither is semi-preferable to the other.

Many simple imperatives might be like this, which partially explains why imperative intensity has received so little treatment in the literature. It's also possible for two imperatives to be *incommensurate* because each is semi-preferable to the other with respect to different third imperatives. Such cases may play a more interesting role in action deliberation when the third imperative is not merely hypothetical but also one that the agent must try to satisfy.

Of course, one must decide which imperative ought to take *priority* in your actions. Priority, as we established previously, may depend in part on facts extrinsic to the imperative—your other plans and desires, the source of the commands, and so on. So priority can be sorted out even in cases where urgencies are *incommensurate*. However, when two imperatives are *commensurate*, then the more intense one ought to be (all things considered) satisfied before, and in preference to, the lesser.

Finally, the present theory can account for cases where intensities are judged on a ratio scale even if they are in fact merely ordered. Suppose $\geq$ has relatively fine-grained rankings over worlds, such that (for example) it distinguishes worlds where you satisfy i and forgo one unit of some good from worlds where you satisfy i and forgo two units of some good. Suppose further we can identify points of indifference where satisfying i is as good as some quantity of x , better than any lesser quantity, and worse than any greater quantity. Given this, suppose we have two imperatives i and j , each of which have this structure relative to the x s, and further that i ranks its satisfaction as on a par with n units of x while j ranks itself as on a par with $2n$ such units. Under such circumstances, we can say that j is not just more intense but *twice as intense as* i .¹¹ One would, of course, need such a structure to be in place consistently—but if it were, it would be natural to speak of ratios of urgencies, not simply ordinal comparisons between them. Note here that the relevant ratio measure has no relation whatsoever to properties, but only to the rankings of different satisfaction and non-satisfaction worlds by the imperative. Hence we can have complex differences in the intensity of an imperative without having to find a sense in which that imperative somehow tracks a magnitude in the world.

3.4 The intensity of pain

The extension to pain is straightforward. Pains are imperatives. As part of their content, they have an especially rich and complex $\geq$, one rich enough to make the sorts of comparisons noted above. So, for example, one pain is more intense than another just in case their world rankings agree that one should tend to the body part involved in the first pain at the expense of the one in the second. Differences in intensity of pains in the same body part can be cashed out (for example) in terms of standard gambles against other valued goods.

This structure explains differences in intensities of pain. It also allows plenty of flexibility to account for oddities of cross-modal judgments of intensity. I can judge the intensity of my pain and my hunger; further, I can judge that my hunger is more intense than my pain. It is not clear to us, however, whether we're ever in a position to judge that some hunger is *twice as intense as* a pain. If that is a puzzling thing to say, the phenomenon can be captured by our account: it will occur if $\geq$ has the requisite structure to make ratio judgments only within modalities, while between-modality imperatives have enough structure to $\geq$ allow ordinal but not ratio judgments. This would be the case if, for example, a set of imperatives had a structure like the following (where $\sim$ indicates 'is ranked equally with'):

- $$\begin{aligned} \geq_1 & \text{ I don't have pain}_1 \sim [\text{I have pain}_1 \text{ and } \$2] > \text{ I don't have pain}_2 \sim [\text{I have} \\ & \text{ pain}_2 \text{ and } \$1] > \text{ I don't have hunger}_1 > \text{ I don't have hunger}_2. \\ \geq_2 & \text{ I don't have pain}_1 > \text{ I don't have pain}_2 > \text{ I don't have hunger}_1 \sim [\text{I have} \\ & \text{ hunger}_1 \text{ and } \$2] > \text{ I don't have hunger}_2 \sim [\text{I have hunger}_2 \text{ and } \$1]. \end{aligned}$$

That is, both rankings place pain only as more intense than hunger, but each of them has a richer structure within just one modality. Hence cross-modality comparisons are merely ordinal, but intra-modality ones can be cardinal.

Further elaborations of the model could account for other pain phenomena. Here is one intriguing possibility. Merely ordinal rankings—even ones that can be used to derive cardinal ranking functions—are not easily aggregated. Even if, say, the pain in my feet is more intense than both my hunger and my thirst, it's not obvious whether I should stop hiking or continue on given that I am *both* hungry and thirsty.

Models of aggregation do exist, however. In those models, different ordinal rankings receive weightings that determine how much they matter to the final aggregate.¹² In the context of pain, those weights might be interpreted as marking individual differences in the importance of pain: individuals who tend to assign low weights to pains in aggregation will, all things considered, discount even relatively intense pains. Hence we might also have a model for individual differences in imperative intensity, one that would further distinguish between different imperative modalities.

Finally—since it can't be emphasized enough—these differences in the intensity of pain depend entirely on the contents of the imperatives that constitute pain. Hence imperativism remains consistent with intentionalism.

3.5 *The naturalization of intensity*

A few final thoughts on the naturalization of imperative intensity will serve as a conclusion. One might worry that the contents so proposed are unusually flrid, particularly the complexities of the ranking function. Note, however, that this is no argument against a model of content itself. One should no more balk at the complexity of the ranking function than at the use of sets of possible worlds. We have not committed ourselves to a story about how the content of imperatives is implemented, but we see no reason why that story faces any principled problems—why the above, for example, could not be cashed out in an approximate sense at least by patterns and frequencies of neural firings.

That said, imperativism is attractive in part because it promises a full naturalization of pain. We conclude, therefore, with two brief reflections on how thinking of pains as imperatives might be especially useful. First, one might look to work on imperatives in natural language, particularly on how the contents of imperatives get fixed in conversational contexts. In this regard we have found particularly illuminating Paul Portner's work on imperatives. According to Portner, when someone utters an imperative in a conversation, it adds an item to the To-do list of the addressee. So, for example, telling me to pass the salt has the effect of appending the property of being such that I pass the salt to my To-do list (2007). The To-do list functions as an action-oriented version of Stalnaker's Common Ground (Stalnaker 2002). It can be affected by other imperatives, as well as by background desires and goals of the participants. A similar model, we believe, might prove fruitful when exploring the effect of pains on the overall mental economy.

Second, imperativism fits naturally with a teleosemantic account of mental content (Millikan 1984; Martínez 2010). The main contention of this kind of approach is that the content of mental states depends on their biological function, or that of appropriately related states. Cashing out that biological function requires a story about the *consumers* of imperative contents. Here again, the imperative view has much to say. Some regions in the motor cortex, for example, are likely consumers for painful mental states. Intensities here play an important role. A mental mechanism can advocate for the spending of more or fewer resources in certain behaviors. Take, for example, physiological responses (increasing heart rate, avoidance) in the face of danger: dangers that are assessed as less pressing provoke a lesser response than more urgent ones. A well-adapted biological agent thus needs both information about what to do and about how important it is to do. Our account of imperative content provides a nice framework within which to cash out that idea.

4 Conclusion

The semantics of imperatives that differ in intensity is not a topic that has received as much attention as others in the study of imperatives. Our discussion of imperative intensity, and the formal apparatus we have used to frame it, is proposed in the spirit of a preliminary exploration. In that spirit, we've shown that imperativism has no insurmountable problems dealing with varying degrees of intensity.

Notes

- 1 See Klein (2007, 2015), Martínez (2015), and Martínez and Klein (2016). Richard Hall has also defended an imperative view in his 2008.
- 2 Recently, it has been suggested that the content of certain perceptual states might have imperatival aspects—see, e.g., Siegel (forthcoming) on *experienced mandates*, or Bengson (in progress) on practical perception. Regardless of whether such suggestions are correct, no one denies that much of the content of perceptual experiences is not imperative.
- 3 We assume this without argument; see Hamblin (1987) for some nice ones. We will refine this quick version, but we will not discuss all of the refinements that might be necessary. Vranas (2008) in particular has given compelling arguments that imperative content requires both satisfaction and violation conditions, an alternative that we won't pursue here.
- 4 A similar, more complex system is still in place. For historical details on the precedence system for Autovon, see page vi of the 1978 "Global Autovon telephone directory", digitized online at <http://hdl.handle.net/2027/mdp.39015078412346>.
- 5 Note that Cutter and Tye use "urgency" rather than intensity. For reasons to become clear shortly, we want to distinguish the two.
- 6 Contrary to what one of the present authors once defended, in Klein (2012).
- 7 Pautz actually raises several objections to imperativism that aren't relevant to the main point of the paper but are worth mentioning. He notes that (i) imperativists have not yet presented a theory of pain locations (true, but a topic for another time; see chapter 7 of Klein, 2015), (ii) that imperativism implies that increasing thermal sensations actually change the type of content they have (true, and here our intuitions

simply differ on how strange this is—after all, the sensation presumably changes for a good reason, namely that the cause has become threatening), and (iii) that imperative theories cannot solve certain problems about magnitude of taste, sound, and color experiences (true, but imperativism doesn't claim that *all* sensations are imperative; these are among those that aren't, and so require a different treatment).

- 8 In that sense, it's even roughly congruent with Pautz's larger project, which is an attack on "tracking externalism" (Pautz 2014). That includes most representationalist theories of pain but arguably does not include imperativism. For a discussion of Pautz's more general argument, as well as a general response on behalf of the representationalist that is also partially applicable here, see Hilbert and Klein (2014).
- 9 Here is one, admittedly very speculative, argument. A true magnitude requires a zero point. While it might seem obvious that there is a zero point for pains—that is, their absence—the idea of a zero point of pain *intensity* is actually quite odd. Intensity is a quality that all pains ought to possess. If the intensity scale has a zero point, then it should be possible to have something that is still a pain, but that has zero intensity. That's a really odd possibility and would need some theoretical motivation. The ordinal version accounts for our intuitions far more naturally: there are no cases of pains with zero intensity, but rather of pains that are simply *absent*. When pains are present, they always have some intensity that's comparable to (i.e. greater than, equal to, or less than) any other pain. Indeed, the imperativist probably ought to say that intensity might be vanishingly mild, but it's never zero—the whole point of imperative sensations is, after all, to motivate you, and a zero-intensity pain wouldn't do that.

Of course, a similar argument could be run for many representational sensations, and there it is more common to assume that they fall on a ratio scale (though this is occasionally debated). Whether this is a reasonable argument would then depend on whether the disanalogy between pain and straightforward representational sensations is strong enough to block the analogy. That is a question for another paper.

- 10 Thanks to Todd Ganson (personal communication) for this suggestion.
- 11 Or more precisely, under conditions where both rankings satisfy the axioms of the von Neumann-Morgenstern theorem. If so, then the theorem would ensure that we could find a suitable cardinal function that described an agent's behavior, and that would in turn be enough to account for ratio judgments of intensity as well as standard bets over merely likely satisfactions of *i*. If you don't like the arbitrary good *x*, you might also construct a cardinal function out of those standard bets. Note that this is consistent with, but does not imply, the presence of an actual real-valued magnitude, which gives rise to $\geq$ (Ellsberg 1954).
- 12 So, for example, one might use the *weighted Kendall-tau distance*. The Kendall-tau distance is a measure of how often reversals occur between pairs of ordinal rankings. Aggregated rankings that minimize the weighted Kendall-tau distance have a number of intuitively appealing properties. See Ailon, Charikar, and Newman (2008) for a discussion and review of other ranking methods, and Martínez (2015) for a ranking-aggregation model of pain.

2 Reasons and theories of sensory affect

Murat Aydede and Matthew Fulkerson

1 Introduction

Some sensory experiences are pleasant, some unpleasant. This is a truism. But understanding what makes these experiences pleasant and unpleasant is not an easy job. Various difficulties and puzzles arise as soon as we start theorizing. There are many philosophical theories on offer that give different accounts of the positive or negative affective valences of sensory experiences. In this paper, we will look at the current state of art in the philosophy of mind, present the main contenders, and then critically compare and contrast them. In particular, we want to examine how they handle the *reason-giving power* of affective states. We will look into two representationalist proposals and a functionalist proposal, and argue that, contrary to their own advertisements, the representationalist proposals don't adequately account for why and how sensory affect can motivate, rationalize, and justify subsequent behavior and intentional mental activity. We will show that our own functionalist proposal does a much better job in this regard, and that when the representationalist proposals are modified to do a better job, they may fare better not because of their representationalist credentials but due to their functionalist ones.

2 Preliminary taxonomy

We will concentrate on those sensory episodes that are unpleasant (including the painfulness of sensory pains)—although most of what we say, when modified appropriately, will be applicable to positive, pleasant sensory episodes. Arguably, some pains—those of pain asymbolics, for instance—are not painful or unpleasant. While we do not discuss such states here, our unifying framework offers a robust means of explaining away such unusual cases.

Our target then are sensory pains understood as sensory experiences that are unpleasant. What is unpleasant about them seems initially to be the experiences themselves. Keeping this in mind, we have the following truism:

(T) Sensory pains are unpleasant experiences.

The job, then, is to specify the following schema as best as one can:

(SCHEMA) Sensory experiences (e.g., pains) are unpleasant if, and only if, ...

Depending on how the ellipsis is filled, theories can be reductive or non-reductive. So, for instance, consider the class of popular *felt-quality* views of sensory pain:

Felt-quality: Sensory (pain) experiences are unpleasant (painful) if, and only if, they have a common phenomenological aspect distinctive of unpleasant (pain) experiences.

As stated, this is a non-reductive view that can be generalized to cover all affective valence—positive or negative. There are different versions of it in the literature, including *hedonic-tone* views and *distinctive-feeling* views.¹ Typically, a defender of these views resists reductive proposals by claiming that the phenomenological aspect in question is a primitive occurrence—either a distinctive phenomenal feature common to all pleasant or unpleasant sensations, or a somewhat varying phenomenal occurrence in a hedonic dimension that cannot be further analyzed. We will not consider such non-reductive views in what follows (but see Aydede 2014).

In principle, a felt-quality theorist could, of course, go on to make a reductive claim by identifying the phenomenological aspect peculiar to sensory affect with a certain sort of neurophysiological event, or with a certain sort of causal role that the sensory experience plays in the larger mental economy of the subject, or with a certain kind of representation or representational content, etc. It may be that these various reductive options remove the appeal of the felt-quality views. Arguably, once a reductive claim is made, all or most of the explanatory work may be done by the reductive component of the theory.

There are two broad approaches one may take in offering a reductive account. First, one could take the task of filling in the ellipsis in SCHEMA to be a matter of supplying a *conceptual truth* derived via the conceptual analysis of pleasure and pain. For instance, the influential *attitudinal* (or, *externalist*) theories developed in value theory and moral psychology can be taken this way. According to such theories, a pleasant sensation is a composite state, composed of a target sensation and some mental (usually, conative) attitude directed toward that sensation, such as desiring, wanting, preferring, or liking the sensation in some sort of way.² The motivation for this view is that the very concept of sensory pleasure or painfulness seems to require that such experiences be essentially connected to other mental states. For instance, one may argue that the very concept of a painful pain that is not disliked or intrinsically motivating is incoherent. Such a state would, according to these theorists, not count as painful at all. Such views are not confined to those in moral psychology; similar proposals have also been made by philosophers of mind, often in the spirit of analytical functionalism. These sorts of views analyze sensory pleasure or painfulness in terms of how these sensations need to be causally related to personal level mental attitudes.³

The second reductive approach one could take is to complete SCHEMA in such a way that the resulting proposal would be *a posteriori*. Rather than filling in the ellipsis merely on the basis of conceptual analysis, we investigate actual states of pain and pleasure and discover their underlying nature. This investigation can involve some speculative theory construction, but it is intended to be scientifically informed and sensitive to empirical facts. In the remainder of this paper we will focus on three such proposals. Two of these theories are broadly representationalist: *evaluativism* and *imperativism*. Both attempt to reduce sensory affect to a certain kind of content (as the names imply, to evaluative and imperative contents, respectively). The third alternative is the account we defend: *psychofunctionalism*. Rather than reduce sensory affect to a certain kind of content, our view identifies sensory affect (painfulness, pleasantness) with the causal and computational roles typical of sensory affect.

Before we consider these views in more detail, we would like to mention briefly another way of carving out the classificatory scheme of accounts of sensory affect.

We started with the claim that (T) is a truism. What makes (T) a truism is that it seems to be an obvious *phenomenological* truth. Anyone who has the relevant concepts and is capable of making mundane introspective observations could, it seems, confirm it instantly without any experimental procedure. Here we understand “phenomenology” to be the relatively robust phenomenology of sensory experiences for which the question of what it’s like to have this or that sensory experience has (usually) a fairly straightforward personal-level answer—even though it may be very difficult or even impossible to express it in public languages. Indeed, to some extent, felt-quality views may be taken as expressing the default common sense view that sensory affect is phenomenologically real and robust.

Many have contested felt-quality views on precisely this point: critics have argued that (T), if it’s a truth, is not a *phenomenological* truth in any robust sense. Indeed, many defenders of attitudinal views (especially in value theory) take their view to be in competition with felt-quality views and do so in a way that denies (T) as expressing an introspectively available phenomenological truth. For the most part, they are motivated by what is sometimes known as the *heterogeneity problem*—the problem of finding a single phenomenal quality common to all and only pleasant sensations (or unpleasant sensations, as the case may be). Once this problem is made clear, (T) as a phenomenological claim seems in jeopardy, and it seems more plausible that something else—an attitude, say—is the common factor between headaches, sun burns, and cases of nausea.

This suggests another possible taxonomy that contrasts *phenomenal realism* with *phenomenal irrealism* about sensory affect. The debate between these groups is an interesting one, for sure—the phenomenology of sensory affect is poorly understood and is certainly in need of better study, philosophical or otherwise. As things stand now, however, it is difficult to settle the debate stated in these terms, as the arguments on both sides tend to rely more or less exclusively on intuition and introspective evidence—the very elements on which the sides disagree. This much is clear: if sensory affect has a phenomenological reality, it is not as robust

as the sensory phenomenology of standard cases of perceptual contact with the extra-mental world.⁴ At any rate, following common sense and the majority of thinkers in this area, we will assume (at least, a weak version of) phenomenological realism about sensory affect for the rest of this paper and read (T) as stating a phenomenological truth. However, nothing much depends on this for what follows—the discussion could be conducted just as well without much modification under the assumption of irrealism.

3 Reductive theories of sensory affect

Let us now consider our three contemporary, *a posteriori*, reductive theories of sensory affect.

We will start with the two representational accounts. Representationalism (sometimes *intentionalism*) is a popular view in contemporary philosophy of mind identifying the phenomenal character of a conscious experience with its representational content. In ordinary cases of perception, the idea is that the phenomenal character of perceptual experience seems exhausted by the features represented in the experience. As Harman says in a classic passage, “When Eloise sees a tree before her, the colors she experiences are all experienced as features of the tree and its surroundings. None of them are experienced as intrinsic features of her experience. Nor does she experience any features of anything as intrinsic features of her experience” (Harman 1990: 39).⁵

Such standard representational accounts of perceptual experience, while popular, have not been without their critics. And, whatever their general problems, it should be clear that, when it comes to accounting for sensory affect, they have a particular difficulty. For, whereas the sensible features represented in a typical visual experience have plausible objective, physical analogues (colors, shapes, locations, distances, textures, etc.), it’s not at all clear what sensible features affective qualities might turn out to be or correspond to. The difficulty with standard descriptive proposals—that is, treating pain as precisely analogous with vision and other perceptual experiences—is that when Eloise feels pain it’s both very difficult to say what external quality she might be representing and to deny that she is aware of some intrinsic, hurtful quality to her experience.

Rather than pursue this quite general debate here, we are going to look at two particular representationalist accounts that seek to avoid the problems faced by unsophisticated versions of the view.⁶ Both seek to carve out a new area of logical space by denying continuity between typical perceptual experiences and cases of sensory affect. Instead, they argue that affective and perceptual experiences share a broadly representational structure (both can be reductively explained by appeal to content), but they critically differ in the kind of content they carry.

3.1 *The imperative view*

The first such view we will consider is imperativism. This is the view that pains and other motivating states specify commands: they involve sensory content that

tells us what to do (or what not to do). Rather than descriptive content, the imperativist believes affective states possess *imperative contents* expressed by these commands. The content constitutive of affective experience is a command to the subject to perform some particular action (or in some cases, to refrain from acting). It is important to understand that, in context, the imperative view arises as a way of saving representationalism (indeed, of saving a strong version of the view) by expanding the possible *moods* that representational content can possess, thereby sidestepping the problems typical descriptive views have with affective states.

The commands adverted to by the imperative view are not taken to be a forced movement of our bodies (it's not a reflex action—as its defenders make clear). Nor is the command specified or explained by a separate attitude like desire, emotion, or belief that accompanies the sensory experience. According to the imperativist, the command is understood as the *imperative representational content* of our affective experiences, and this content in imperative mood is the reductive explanatory base for what it's like to have an affective sensory experience of that kind.⁷

Richard Hall (2008) argues that itches (but also hunger, thirst, and pain) are best understood by their possession of imperative contents:

On my view, itches are experiences with a content, as the intentionalists argue. But the intentional content isn't a representation of some bodily state ... 'Representation' suggests descriptive content but itch experiences don't have descriptive content. Itches don't describe some state of the body at the felt location. Rather the intentional content of an itch is an imperative: 'Scratch!' Not in English, of course, but in sensory code.

(2008: 526)

On Hall's view, the imperative content is the motivating and affective part of the experience and can be understood as part of an overall experience that also contains (some) descriptive contents.

Colin Klein (2007) defends a more austere version of the view, this time explicitly directed at pain:

I argue that pains are exhausted by their content, but that this content is imperative rather than representational. Pains thus command rather than describe. Commanding is still a way of having content, however, and so intentionalism is preserved.

(2007: 518)

Klein's view explicitly parallels strong versions of the representational view, only swapping out imperative content for descriptive. Instead of positive imperatives like "scratch!", "drink!", and "eat!", Klein argues that pains involve negative imperatives: "Do not <Action>!" For example, a pain in the foot might have the content: "Do not place weight on the foot!"

More recently, Manolo Martínez (2011) has also defended a version of imperativism for pain. According to Martínez, “The content of a headache is one with satisfaction conditions that we could render more or less as: *that this bodily disturbance is no more!*” (Martínez 2011: 78). He goes on to suggest some advantages this particular formulation has over Hall’s and Klein’s versions, and to suggest some ways that the view can be naturalized (more on this last point shortly).

While these versions differ, sometimes importantly, in their details, for our purposes these differences don’t matter. What is crucial for us is that all versions of the view appeal to a special representational content that is understood in terms of commands. Generalizing the view:

Imperativism: S’s pain is unpleasant if, and only if, S receives a command from her relevant sensory system to the effect that S engage (or, stop engaging) in a certain sort of behavior directed to the bodily location where the pain is felt.

A few clarifications are in order before we can assess the general view. The imperativist needs to specify who or what issues the command as well as who receives it. Issued commands are supposed to be representational items with satisfaction conditions. These conditions need to be specified at least in general terms. Similarly, commanding requires a representational medium in which commands are issued. Thus, the format of commands needs to be specified too.

We will assume (as Hall makes explicit) that the representational medium is a sensory code in which our sensory experiences are realized. So the representation in question is experiential (perhaps, non-conceptual with presumed phenomenology) and issued in the imperative mood commanding the subject to do whatever it takes to change the disordered bodily condition that is the cause of the pain. We will follow Martínez (2011, 2013) in taking this command to be something like:

Pain command: See to it that this bodily disturbance is no more!

The command is about a physical disturbance in or on the subject’s body to the effect that the subject take steps to get rid of it. Let’s also say that the command is issued by the subject’s pain module—a specialized sensory system that monitors and reports on potential bodily disturbances or disorders. This command is issued to the subject and is received when the subject experiences the pain as unpleasant or painful. These glosses will suffice for the moment.

The key question concerning imperative contents for pain concerns how such contents gain their imperative mood. As Hall notes in a footnote:

How can a sensory experience have an imperative mood or force? Although possible answers come to mind (perhaps Mentalese, or more particularly sensory-ese, contains explicit indicators of imperativeness, or perhaps the functional role of the message—where it’s from, where it goes, how it’s operated on—constitutes its imperativeness), this is a topic for another paper, since it

is a general problem for any intentional view of experience. How, after all, does a ‘descriptive’ representation, say from vision, get its indicative mood and its declarative or assertive force?

(2008: 526)

Note that Hall here openly grants the possibility that the functional role of a representation is essential for its content to be imperative.⁸ This becomes clear when he suggests that this is not a worry since the indicative versions of representationalism face the same problem. This is somewhat misleading, however, for the most plausible accounts of how to cash out indicative contents—various tracking theories of representationalism—are simply not available for typical motivating states (see Aydede and Fulkerson 2014). We will argue below that this problem is not so easily dismissed, and that the only plausible way to cash out imperative content is by invoking some version of (psycho)functionalism.

3.2 *Evaluativism*

Let’s now turn to the other form of representationalism, a view that David Bain calls *evaluativism*.⁹

According to the evaluativist, affective sensory states can be reductively explained by their *evaluative content*. Such content is not purely descriptive (as in ordinary perceptual representational views), nor are they commands as on the imperative view. Instead, the view holds that affective experiences involve contents that make some normative or evaluative claim on the world, usually in terms of something being good or bad for the subject. Again, generalizing away from the details of the various extant proposals (and following Bain’s canonical formulation of the view), we can state the essential elements of the view:

Evaluativism: “A subject’s being in unpleasant pain consists in his (i) undergoing an experience (the pain) that represents a disturbance of a certain sort, and (ii) that same experience additionally representing the disturbance as bad for him in the bodily sense.”

(Bain 2013: S82)

Again, these representations are experiential and (perhaps) non-conceptual with appropriate phenomenology. It is the second clause that makes this form of representationalism evaluativist. The claim here, unlike that made by the imperativist, is that the mood of this representation is, like those for normal perceptual experiences, indicative. Or at least, it is a representation that is truth-apt, with legitimate truth-conditions. What differs from the standard perception case isn’t the *mood* of the content, but rather the nature of the properties represented. Whereas standard perceptual experience puts us into contact with neutral sensible qualities, affective experiences put us into contact with value-laden or normative qualities. These are still thought to be objective features on most versions of the view (although there is interesting logical space here for a non-naturalist evaluative

theory, including a so-called *Edenic-account*—see Chalmers 2006) characterized by what is, as a matter of fact, good for the subject (or, on some accounts, the subject's body). These contents are normative in the sense that they specify, for example, what qualities an agent *should* or *should not* seek to avoid, or by stating which qualities are intrinsically *good* and which *bad*. The main challenge for such a view, as one might expect, is to give a plausible account of the nature of these evaluative contents. Various proposals have been made in the literature, and some of them (i.e., Cutter & Tye 2011) are fairly complex. Our view is that none of the proposals so far introduced have given an adequate, naturalistically satisfying account of the nature of evaluative content. (We should be clear, this does not necessarily stem from an aversion to all normative content, only to the use of such content to explain episodes of affective sensory awareness).

3.3 Psychofunctionalism

Finally we turn to a rough sketch of our preferred psychofunctionalism. Unlike the previous accounts, the psychofunctionalist view does not attempt to reduce the phenomenal character of sensory affect to a special kind of representational *content* or to a special *mood* for such content (this is not to say that the view denies pleasant and unpleasant states involve representations—see below). Instead, the psychofunctionalist can trace their motivation to the idea present in analytic functionalism that pains and other affective sensory states seem to be essentially linked to a suite of behavioral and dispositional reactions. Unlike these earlier forms of functionalism, the psychofunctionalist is not interested in mere conceptual analysis or in giving an account in behaviors and dispositions characterized solely at the personal level. Instead, the view works on the assumption that our best account of the causal influences, computations, and dispositions constitutive of the target mental state will be uncovered in (largely) subpersonal terms by work in the relevant empirical sciences (especially neuroscience and cognitive psychology). The reductive base then may not include causal roles immediately available to introspection or to our naïve conception of the ordinary notions involved. The causal roles involved will be, for the most part, a substantive and theoretically robust *discovery*, couched ultimately in the language of the most reliable empirical sciences.

For our purposes, we can, to some extent, stay neutral on the nature of the functional roles to be uncovered. While any specific view will ultimately have to make some choices here, our defense of the view can also remain silent about more general issues concerning whether the functional roles are purely causal, computational, or teleological; or whether the ontology of functions is process-based or substantival (or both). Indeed, we want to leave open the possibility that psychofunctional roles may be constituted by distinct kinds of functions. Generalizing the view:

Psychofunctionalism: A subject's pain experience is unpleasant if, and only if, the sensory information carried by the pain experience plays a certain causal/psychofunctional role (call this role CR).¹⁰

As we noted, specifying CR isn't exclusively or even mostly a philosopher's job. Certainly not at the most detailed levels. Nevertheless, it is important to fill in the ellipsis by outlining the kind of causal role we have in mind. Work in recent affective neuroscience suggests that this role consists of a complex processing of incoming sensory information that, among other things,¹¹

- sets interruptible motivational parameters (motivational biasing—"more-of-this" or "less-of-this" or "stop-this" incoming stimulus);
- prepares and primes the effector or psychomotor systems of the organism, providing action-preparedness (motor biasing);
- provides appraisals of the incoming sensory information for its significance for the organism and for its potential for enhancing its behavioral repertoire (epistemic biasing);
- influences action preferences on the basis of the sensory stimuli's informational content for present and future behavior through associative or cognitive learning, habituation, incentive sensitization, etc.; and
- provides steady, earmarked input to more centralized concept-wielding cognitive and decision-theoretic or conative systems.

One crucial aspect of this proposal is that the distinctive affective processing is inherently motivational in a broad sense. This leads us to call this style of processing *m-processing*. Keeping in mind that while m-processing is a general term for a heterogeneous class of distinct functional processes and realizers,¹² we can still state a non-trivial explanatory reduction in terms of a theoretical identity by holding that sensory affect *just is* the m-processing of incoming sensory signals. We will say that the affective valence of sensory experiences (including the unpleasantness of pains) consists of the sensory information contained in the experience being m-processed, i.e., being subjected to the kind of processing outlined above and whose details are to be filled by empirical sciences. We believe that the metaphysics of sensory affect is a functional affair and that this is in fact supported both by findings in the sciences but also (to some extent) by the phenomenology of such experiences. Let us elaborate on this a little further.

The functional nature of m-processing is not altogether hidden at the personal level—this is what makes attitudinal theories, analytic functionalism, and in general conceptual reductionism, *prima facie*, plausible. The pleasantness of the taste of a ripe strawberry consists literally in our being motivationally (motorically, epistemically, etc.) *biased* toward that very taste *and* toward the strawberry in the very experiencing of that taste. This is how that taste *and* the strawberry are being presented to our consciousness. The pleasantness or the unpleasantness, in this sense, is the "felt evaluation" of that taste and the strawberry (Helm 2002). Clearly, some sort of a "desire-like" state is involved somewhere in the phenomenology of sensory affect. When we have a pleasant sensation such as the taste of the ripe sweet strawberry, there is some sort of a tug that we feel toward the taste, we feel some kind of pull that makes us "want to continue having the

sensation.”¹³ Not only that, and more importantly, the pull is also toward the strawberry, the object of our taste experience. But the ordinary intuitive notion of desire isn’t up to capturing this. If there is any sense in which there is a kind of desire involved in affective experiences, it is a *phenomenologically salient experiential desire*. Let us call this *phen-desire*. Just as the sensory informational content of a perceptual experience can be encoded non-conceptually in the sensory code (descriptive sensory phenomenology, call it *phen-believing*),¹⁴ the affective content of the experience can be identified with the way this sensory information is m-processed for motivational/learning related purposes (affective/conative phenomenology).¹⁵ It is a form of *experientially desiring* (phen-desiring) the object of the experience whose sensible properties are also phenomenologically registered (phen-believed) in the sensory content of the very same experience. So we have the *experiential* (non-conceptual) analogs of the ordinary notions of belief and desire (these latter generally conceived as conceptually structured propositional attitudes).

In the case of normal sensory pains, the experience of pain would involve sensorially encoding the physical disorders in the body while presenting them as the objects of our phen-desires (maybe to the effect that they cease or go away). The phenomenology of this phen-desiring would be tantamount to the pain’s being painful, unpleasant—its negative affective character. We would like to note here that the affective qualities such as pleasantness and unpleasantness (the affective phenomenology) primarily qualify the affective *experiences* as loci of information processing, not the things or events this information is about. Phen-desiring may be a sensory desire for the extra-mental objects the sensory information is about, but the phenomenology is primarily the phenomenology of the phen-desiring ($\approx$ m-processing), not of its object. What is pleasant or unpleasant are the experiences, not the objects of those experiences. If there is a sense in which the *objects* of our sensory/affective experiences are pleasant or unpleasant, it is only in a derivative sense: they are apt to cause pleasant or unpleasant experiences. The primary locus of affect is the experience itself.¹⁶ It is important to keep this in mind because it will be crucial in understanding how the experiences themselves can be (intrinsically) good or bad (as opposed to their objects being so) and can thus be the objects of our ordinary desires—we will return to this below.

Thus, even though the metaphysics of sensory affect is psychofunctionalist ($\approx$ negative or positive m-processing), our view has a lot of insights to offer at the personal (phenomenological) level by identifying this m-processing with (roughly) a kind of pro or con experiential desire (phen-desire) that can make the experiences themselves be the objects of our ordinary desires. (See Figure 2.1.)

All of this is moot, however, if it turns out that psychofunctionalism is also the only game in town. In what follows, we argue that, when modified appropriately, the two competing alternative accounts are not, in fact, competitors with the psychofunctionalist view. Instead, they are either gussied up versions of the view when they work, or else they are just inadequate or offer at best metaphors instead of a theory.

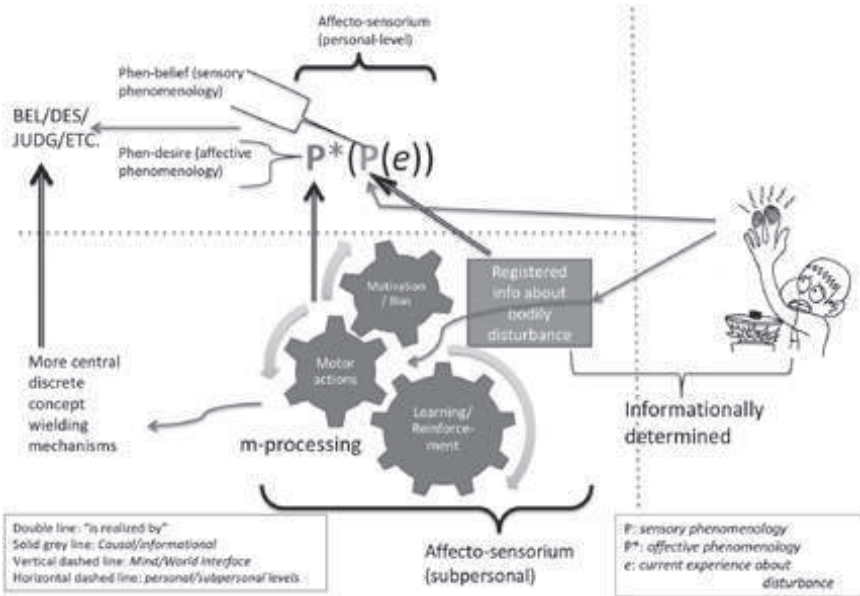


Figure 2.1 Painful pains according to psychofunctionalism—this basic framework can be generalized to all sensory affect, positive or negative.

How to adjudicate between these three reductive theories? There are basically two ways of doing this. One is to compare the theories against well-known, or at least agreed upon, facts holding of the target domain these theories are theories of. These may be observational, theoretical, or normative facts. We can try to find out which theories do a better job in explaining or accommodating these facts. The other method is to describe difficulties internal to each theory. These may relate to the assumptions they make, or their observable consequences, or their formal or theoretical commitments, etc. We can then find out that some of these difficulties make the theories implausible or unlikely, and the attempt to fix them may collapse a theory into another, etc. We have arguments of both kinds against evaluativism and imperativism and for psychofunctionalism. In this paper, however, we will be concerned only with comparative arguments of the first kind, comparing how they fare about the motivating and justificatory reason providing force of sensory affect.

4 Sensory affect and reasons

To simplify the discussion and allow us to focus on the relevant contrasts, let's work on an example. Suppose Sally was making a grilled cheese sandwich and burned her finger accidentally on the hot pan. Let d be the burnt condition of her finger. Let e be the somatosensory experience caused by this in the standard way. Thus, e is Sally's pain experience of her burnt finger, and it is deeply unpleasant

(painful). According to the three theories then, and simplifying to bring out what is essential:¹⁷

Imperativism: e 's unpleasantness $\approx e$'s incorporating an experiential command to Sally, <see to it that d is no more!>.

Evaluativism: e 's unpleasantness $\approx e$'s experientially representing d to Sally as bad.

Psychofunctionalism: e 's unpleasantness $\approx$ m-processing of the sensory information (about d) contained in e . (Roughly: phen-believing that d is occurring and phen-desiring that d is no more—see below.)

One way to evaluate these competing theories is to look at how well they explain certain truths about painful pains (and other affective experiences). Two of these (used by Bain himself against the imperativist and attitudinal views) are of special importance precisely because they are thought to pose particularly hard problems for non-representationalist theories. These are the motivational (rationalizing) and justificatory roles of painful episodes (what Bain 2013 calls their “hedomotive” role). Pains not only motivate but also normatively justify certain kinds of behavior as well as the formation of certain beliefs and desires. In other words, they provide motivating and justifying (good) *reasons*—albeit defeasible reasons—for intentional action. Let us explain.

Consider first pain's motivating role. When Sally burns her finger, she doesn't simply experience the pain, e . In addition, and *because of* e , she engages in complex intentional actions and forms new intentional states downstream. When we ask what motivated Sally's actions, we want to know *why* Sally performed all these complex actions. Consider what actions Sally does after she burns her finger: she quickly moves her finger away from the pan, she screams and immediately seeks out ice and cold water for her hand, a bit later she puts soothing and anti-inflammatory cream on her finger. Note that these are pieces of behavior directed toward the condition of her finger, d . Call this complex set of behavior *d-directed* behavior.¹⁸ But this is not all that happens: her intentional mental states also undergo complex changes *because of* e : her attention is drawn immediately to the injury, the experience generates stress and concern, and produces in her a set of beliefs and desires:¹⁹

- (d-Belief_i) a belief that she burned her finger (i.e., d occurred),²⁰ and
- (d-Belief_n) that this is bad for her (i.e., d is bad for her), and
- (d-Desire) a desire to get rid of d (i.e., that d cease or be no more).

Sally also has strong experientially directed beliefs and desires:

- (e-Desire) the desire that the pain experience, e , itself stop, and
- (e-Belief_n) a belief that e is bad for her,

and as a result, she also engages in behavior directed to stopping the pain itself—for instance, by taking pain killers—call this *e-directed* behavior.

Sally's pain experience not only *causes* her *d*- and *e*-directed behaviors but also *rationalizes* them in the sense that it makes them intelligible and rational from her subjective perspective. Given her experience, it makes rational sense to engage in these kinds of behavior. So when we say that *e* motivates Sally's behavior, we don't have mere causation in mind but also rationalization of action.²¹ Sally's pain experience provides reasons *for which* she acts the way she does.

Furthermore, these are *good* reasons that rightly *justify* Sally's actions of both kinds. When we ask what *justifies* Sally's actions, we are not seeking a merely motivational/rationalizing story from her own perspective. Instead, we want to know if the actions were performed for *good* reasons.²² The actions that Sally performs are justified by her painful experience and her burnt finger. In other words, her painful experience and what it indicates give her perfectly *good* reasons to seek ice, avoid the hot pan, *and* take pain killers.

We have, then, the following claims ordinarily believed to be expressing self-evident truths:

- (**d-Motivation**) *e*'s painfulness *inherently* (but defeasibly) motivates Sally's *d*-directed behavior, which is to say that it defeasibly motivates it all by itself without the need of any beliefs and desires.²³
- (**d-Justification**) *e*'s painfulness, insofar as *d* is in fact bad for her, *provides* Sally with a good reason for her *d*-directed behavior—it *justifies* her *d*-directed behavior (as well as her belief that *d* is bad for her d-Belief_n).
- (**e-Motivation**) *e*'s painfulness *inherently* (but defeasibly) motivates Sally's *e*-directed behavior, which is to say that it defeasibly motivates it all by itself without the need of any beliefs and desires.
- (**e-Justification**) *e*'s painfulness is *intrinsically* bad, which is to say that it *constitutes* a good non-instrumental (i.e., justifying) reason for both Sally's *e*-directed behavior, and her *e*-Desire and *e*-Belief_n.²⁴

It is certainly not out of the question to deny these claims as expressing truths. Indeed, Klein (forthcoming) seems tempted to deny both (d-Motivation) and (e-Motivation). Cutter and Tye (2014) flat-out deny both (e-Motivation) and (e-Justification).²⁵ But given a modest phenomenological realism about sensory affect, we will not question these claims. Our concern is to find out which of the three accounts does a better job of explaining why they hold.

4.1 Evaluativism and reasons

One of the major motivations for representationalist views of affect is the thought that no state can rationalize or justify behavior and the formation of other intentional states without possessing a representational content. So, for instance, a major complaint against felt-quality views of sensory affect is that the occurrence of a mere phenomenal character or a quale cannot inherently rationalize or intrinsically justify—a felt-quality is simply not the right kind of thing for that. So the evaluativists think that their account is best suited to explain

why the above claims hold. Affect can motivate/rationalize *and* justify because it is the experiential representation of an evaluative state of affairs. If we replace “e’s unpleasantness” with “e’s representing *d* as bad” in (d-Motivation) through (e-Justification), we obtain the evaluativist formulations of the claims that are supposed to remain true. Let’s start with

(d-Motivation_{eval}) *e*’s representing *d* as bad inherently (but defeasibly) motivates Sally’s *d*-directed behavior, which is to say that it defeasibly motivates it all by itself without the need of any beliefs and desires.

This seems to be a version of motivational internalism—the view that representing evaluative/normative state-of-affairs in a truth-apt way has an internal connection to motivation. This is a controversial thesis, but we do not want to rely solely on this for our criticism since we will offer a more direct criticism shortly. We are officially neutral about motivational internalism. For the moment, we simply note that evaluativism seems to require the truth of this controversial thesis (cf. Cohen & Fulkerson 2014).

What about (d-Justification)?

(d-Justification_{eval}) *e*’s representing *d* as bad, insofar as *d* is in fact bad for her, provides Sally with a good reason for her *d*-directed behavior—it *justifies* her *d*-directed behavior (as well as her belief that *d* is bad for her, d-Belief_n).

This seems correct as far as it goes. But how far does it go? We need to know what it is for *d* to be in fact bad for Sally. Suppose that Sally starts to develop allodynia around the location of *d* (i.e., even innocuous non-nociceptive stimuli around *d* start to cause painful pains in Sally). So for instance gently touching her palm or the back of her hand or wrist generates painful pains.²⁶ Gently touching her palm (call this *g*) is not in any physical or objective bodily sense (i.e., apart from causing a pain in her) bad for her. Sally now has an experience, *e**, that misrepresents *g* as bad. This still makes her experience, *e**, painful/unpleasant, to be sure. But what becomes of (d-Motivation) and (d-Justification), when reformulated with this situation in mind?:

(d-Motivation*) *e**’s representing *g* as bad inherently (but defeasibly) motivates Sally’s *g*-directed behavior, which is to say that it defeasibly motivates it all by itself without the need of any beliefs and desires.

(d-Justification*) *e**’s representing *g* as bad, insofar as *g* is in fact bad for her, provides Sally with a good reason for her *g*-directed behavior—it *justifies* her *g*-directed behavior (as well as her belief that *g* is bad for her, *g*-Belief_n, below).

Sally may not know that *g* is not bad for her (in the physical sense), so her *g*-directed behavior and propositional attitudes may well be inherently motivated and rationalized from her own subjective perspective. Furthermore, because *g* is

not in fact physically bad, it doesn't seem unintuitive to claim that Sally doesn't have a good non-instrumental reason for her *g*-directed behavior. So far so good.

But suppose that she *does* know that *g* is not physically bad for her. Her experience, *e**, will of course continue to represent *g* (incorrectly) as bad. Let's have a look at her belief/desire set under this scenario. She now forms:

- (*g*-Belief_{*f*}) a belief that *g* is occurring
(i.e., a point in her palm is being gently touched), and
- (*g*-Belief_{*n*}) that *g* is *not* bad for her, and
- (*g*-Desire) a desire to get rid of *g* (i.e., that the gentle touch stop).

Sally also has a strong desire

- (*e*-Desire*) that the pain experience, *e**, itself stop, and
- (*e*-Belief_{*n*}*) a belief that *e** is bad for her.

Note that *g*-Desire is now *not* rational (even from her own subjective perspective) unless we provide a different rationalization from Sally's perspective. A rationalizing belief is not difficult to find:

- (*g*-Belief_{*n-r*}) that *g* is bad for her (in the sense that *g* is causing *e**, which is, and is believed by Sally to be, bad).

We can mark this relational sense of "bad" (causing painful pains) with "r-bad." Note that the sense in which painful pain experiences are themselves bad, as *e*-Belief_{*n*} and *e*-Belief_{*n*}* express, is still different. Painful pains are bad not because they are themselves causing painful pains. Presumably they are also not physically bad—at least not in any sense in which their objects, physical disorders, are physically bad. We may capture this third sense by saying that the painful pains or pleasant sensations are *intrinsically* bad or good. We will say more about this below.

But now note the following. This second scenario suggests that the rationalization and justification of Sally's desire to get rid of *d* (*d*-Desire) and the subsequent *d*-avoidance behavior, in the original scenario, has in fact very little to do with the evaluative content of her experience, *e*, that *d* is bad in the physical or some objective bodily sense. Because the second scenario suggests that even if *e* misrepresented *d* as bad, Sally's desire (*d*-Desire) and her *d*-avoidance behavior would still be rational and justified as long as getting rid of *d* is a way of getting rid of her painful pain.²⁷ This raises the question of: "Which belief is rationalized and justified by Sally's pain experience in the original scenario where she burns her finger?" The belief

- (*d*-Belief_{*n*}) that *d* is bad for her (in the physical sense),

or the belief

(d-Belief_{n-r}) that *d* is r-bad for her (i.e., *d* is bad in the sense of causing a bad experience)?

Intuitively the right answer is: both—although d-Belief_n may now have less direct justification (see below). But the evaluativist cannot agree that d-Belief_{n-r} is rationalized and justified by the representational content of *e*. For the evaluative content of *e* is not the content of d-Belief_{n-r}. The evaluativist needs to explain why painful pain experiences, i.e., experiences that merely represent a body part to be physically bad, are themselves bad, so that much of our body-directed desires and avoidance behavior would remain the same (rationally and justifiably) even when our painful pains misrepresent. This the evaluativist cannot do by merely adverting to the existence of evaluative content. Here is another way to bring this out.

Suppose Sally doesn't have any opinion about whether she has developed allodynia—indeed, she doesn't even know what allodynia is. For all she knows, gentle touch may be causing some physical damage (given the slight swallowing and reddening of the over-sensitive area), or maybe not. Still, when slightly touched, she feels pain, and forms the belief that

(B) *that* is bad,

where “that” refers to the gentle touching, *g*. We assume that Sally knows what she means and is intuitively justified in her belief by her experience. But Sally is surely *not* prepared to withdraw her judgment (B) if she learns that *g* is not damaging or bad in any physical or objective bodily sense. Just tell her, “But Sally, gentle touching is not in any physical sense bad, damaging, or disturbing, you are just wrong in thinking *that* is bad for you.” The most likely response you will get is: “It is bad, it still hurts!” But this just means in her mouth that *g* is causing a painful pain in her. This seems to show that Sally has in mind (g-Belief_{n-r}) from the beginning—in other words, even if, contrary to imagined facts, *g* turned out to be physically bad/damaging.

However, if epistemic justification is (at least, partly) a matter of representational content, and evaluativists are right about the content of pain experiences, Sally's experience, even if it may rationalize it, does not justify (B)—so none of her *g*-directed subsequent behavior is justified either—if (B) attributes physical badness/damage to *g*. For (B) is just false according to evaluativism under the imagined scenario (just as the evaluative content of her pain is), despite Sally's protests to the contrary.²⁸

It may be that in normal cases, both the experientially represented damage and its experiential representation (pain) are bad. But evaluativism seems incapable of explaining why our beliefs and desires subsequent to our painful pain experiences seem to be aligned in their content with the intrinsic badness of experiences and the r-badness of their causes, rather than the physical badness of the damage. So, contrary to their own advertisements, evaluativists do not have an adequate account of how painful pain experiences provide justifying reasons.

This brings us to the other two truths of the four stated above. According to evaluativism, the proper readings after the required substitution are:

- (**e-Motivation**_{eval}) *e*'s representing *d* as bad *inherently* (but defeasibly) motivates Sally's *e*-directed behavior, which is to say that it defeasibly motivates it all by itself without the need of any beliefs and desires.
- (**e-Justification**_{eval}) *e*'s representing *d* as bad is *intrinsically* bad, which is to say that it *constitutes* a good non-instrumental (i.e., justifying) reason for both Sally's *e*-directed behavior and her anti-*e* desire (e-Desire) and belief (e-Belief_n).

Far from being self-evident, these do not seem to be true at all.²⁹ Philosophically, one of the most interesting and important things to be explained about sensory affect is why the affective experiences themselves are pleasant or unpleasant, and therefore, intuitively, are non-instrumentally good or bad to have. It seems to us that evaluativism does not touch this issue at all. Focusing on the case of pain sometimes obscures this point, since normally painful pains signal the objectively bad things happening to parts of the body. Because of this, one's intuitions do not usually settle the claim that if it weren't for the badness (unpleasantness) of pains, what they indicate wouldn't be bad or unpleasant. Consider the intuition that if it weren't for the goodness or pleasantness of pleasant sensations, the objects of these sensations (e.g., the strawberry or its taste objectively understood) wouldn't be pleasant or good. It is clear that the taste of a strawberry (objectively understood), or the strawberry, for that matter, is good or pleasant only in the sense that it is pleasure-causing, i.e., what is good or pleasant in the first instance is the experience of pleasure—the pleasantness of the taste sensation. Similarly, in the pain case: what is important from the point of view of understanding our pain-based beliefs and actions is the badness of painful pains, not the badness—or represented badness—of the damage or other stimulus.³⁰ It seems to us that the intuitions about the *primacy* of affect as attaching to the experiences/sensations are fairly secure (see our 2014 for more elaboration). Saying that this experiential affect is merely the representation of their objects as good (objectively or physically), or as pleasant, for that matter, does not take us far.³¹

David Bain (2013) briefly responds to the worry about how the possession of *d*-directed evaluative content can make the experience itself intrinsically bad and can therefore motivate *e*-directed behavior. As we argued above, this is particularly worrisome when the evaluative content is false and known to be false by the sufferer. As Bain puts it: "Why shoot the messenger if you know the message is false?" (2013: S86).³² He points out the existence of other instances where it seems very natural to think of these states being themselves bad for the subject. He claims that we have an intuitive understanding of why, for instance, being in emotional states such as grief or fear is itself intrinsically bad for the experiencer. These are cases in which something *experientially* strikes the subject as bad, and because of this, Bain seems to claim, being in such states is itself intrinsically bad for the subject.

But even if we accept the analogy with emotions (which is moot), it only shows that there are other cases where we may be paying more attention to the messenger than to the message. In fact, we don't need to refer to emotions for good examples for this. Pretty much any experience with medium-to-strong negative or positive affective character will do. When I find the taste of strawberry so pleasant, or am having an orgasm, I am very much into the messenger, not (merely) the message (i.e., these are states I have direct interest in because they themselves are intrinsically good or bad to be in). But this is just to repeat the main worry: how can experientially representing bodily badness or goodness itself be intrinsically good or bad? But we thought that an essential part of the philosophical project about sensory affect was to explain just this on evaluativism: how could representing bodily badness or goodness metaphysically amount to experiential unpleasantness (badness) or pleasantness (goodness), and how can it therefore motivate and justify our *e*-directed behavior and propositional attitudes (*e*-Motivation and *e*-Justification, above)? Pointing to even more problematic examples like positive and negative emotions as existence proofs amplifies the mystery rather than answering it.^{33,34}

4.2 Advantages of psychofunctionalism

Note that psychofunctionalism does not suffer from *any* of these problems. According to our view, negative or positive affect is the *m*-processing of incoming sensory information, which is, roughly, the phen-desiring or phen-aversion of what the sensory information is about. Because the incoming information is encoded in sensory code being subject to *m*-processing (with something like CR), pleasant sensory experiences are pleasant because their pleasantness is constituted, *very roughly*, by the satisfaction of phen-desires as registered by the simultaneous phen-beliefs. Unpleasant sensory experiences, including painful pains, are unpleasant because their unpleasantness is constituted, *very roughly*, by the frustration of phen-desires as registered by the relevant phen-beliefs.³⁵

Even given this much, we can explain how the painfulness of Sally's pain, *e*, motivates/rationalizes as well as justifies both Sally's *d*-directed and *e*-directed propositional attitudes as well as behavior. Her *d*-directed propositional attitudes,

(*d*-Belief_i) her belief that she burned her finger (i.e., *d* occurred), and
 (*d*-Belief_n) that this is bad for her (i.e., *d* is bad for her), and
 (*d*-Desire) her desire to get rid of *d* (i.e., that *d* cease or be no more),

are rationalized and justified by her experience, no matter how "bad" is interpreted—although we usually have relational badness in mind. Sally's experience presents *d* (or *g* for that matter) as the object of her frustrated phen-desires. Similarly, Sally's *e*-directed propositional attitudes,

(*e*-Desire) the desire that the pain experience, *e*, itself stop, and
 (*e*-Belief_n) the belief that *e* is bad for her,

are rationalized and justified by her experience, although in a different way. How?

Recall what we said earlier: if there is a desire involved in affective experiences, it is a phenomenologically salient *experiential desire*. We called this a *phen-desire*, and contrasted it with the sensory informational content of a perceptual experience (called *phen-believing*). On this model, the affective phenomenology is one primarily of phen-desiring and its continuing frustration (or satisfaction) as registered by the incoming sensory information (phen-believing); the primary locus of the relevant sensory affect is the experience itself. So it follows that Sally, in having a phen-desire, is undergoing (and aware of) an experience that is intrinsically bad (the negative m-processing of sensory input—roughly: phen-desire frustration—is itself bad). Her forming *e-Desire* and *e-Belief_n* (as well as her introspective belief that *e* is occurring) shows Sally's direct (intuitive) cognitive and conative grasp of this normative fact.³⁶

So when Sally experiences the sharp burning pain, she is completely rational in seeking to end that experience. This is so even if there is no damage at all in her body, or any threat of damage, or the like. The experience itself is negative (unpleasant, intrinsically bad), and one she is both rational and justified in ending.

More importantly, her phen-beliefs and phen-desires also motivate and justify her *d*-directed behaviors. Sally's phen-belief (the descriptive sensory phenomenal content) that *d* is occurring rationalizes, and if correct, justifies *d-Belief_i*. When she burns her finger, she becomes experientially aware that *d* is happening, and so she is rational to act in ways reflecting that awareness. *d-Belief_n* is motivated and justified in virtue of the fact that her experience is constituted by both the phen-belief concerning the disturbance and a frustrated phen-desire about this disturbance. Recall that *d* is simultaneously the object of her phen-desire and phen-belief. (It is the very same sensory signal that is subjected simultaneously to both indicative and conative processes, so to speak.) While the negative phenomenology primarily attaches to the experience constituted by her frustrated phen-desire, her phen-desire itself is directed at the disturbance, *d*. So, her phen-desire (the affective phenomenal content) that *d* be no more both rationalizes and justifies her *d-Desire*. Note that Sally's phen-desire and *d-Desire* occur at different levels (the former is experiential, the latter is a propositional attitude) and thus do not compete with one another as either causes or reasons. It is perfectly natural to say that Sally has her *d-Desire* *because* of her phen-desire, and that the phen-desire *provides* a *good reason* for her *d-Desire*. It is also perfectly natural to say that Sally's frustrated phen-desire *constitutes* a *good reason* for her *e-Belief_n* and *e-Desire*. There is no threat of circularity or emptiness in either of these formulations.³⁷

We now have fairly straightforward experiential equivalents of standard belief-desire explanations (both causal and rationalizing), and, as we should expect, the normative justification goes along with the truth-value of their content. Clearly, if we are right about the rationalization and justification of Sally's relevant propositional attitudes, how to extend this account to cover how both *e*-directed and *d*-directed *behavior* are motivated and justified should be clear. Furthermore,

our account has the additional advantage of showing how the relevant range of behavior can be motivated even in the absence of the *actual* formation of propositional attitudes both about *d* and about *e*—if that is possible.³⁸

Finally, a few words about the metaphysics of intrinsic goodness or badness. We identified sensory affect, pleasantness, and unpleasantness, with the satisfaction and frustration of phen-desires respectively, which are in turn identified with positive or negative m-processing of incoming sensory information. One may reasonably ask: what is it about the frustration of a phen-desire that makes it intrinsically bad? Granted, painful pain experiences are intrinsically bad—maybe obviously so. But it is not obvious that phen-desire frustration that we claim makes pains painful (unpleasant) is intrinsically bad. This is a legitimate worry for those inclined towards non-reductionist or non-naturalist accounts of any sort of intrinsic value. Sensory affect, as we understand it, is a species of intrinsic value—we are officially neutral about whether there are other species and whether they can be reduced or naturalized if there are.³⁹ But we do want to claim that the intrinsic badness or goodness of a sensation (qua sensation) metaphysically consists of phen-desire frustration or satisfaction as registered by the relevant simultaneous phen-beliefs. The intuition here is that *if* sensory affect's having intrinsic value is not an ontologically primitive occurrence, *then* our proposal is the best reductive account available for its metaphysical nature. The possession of intrinsic value by sensory affect, we claim, is made more intelligible by our reductive account. We cannot defend this claim here. But, as naturalists, we find it intuitively plausible and illuminating.

4.3 Imperativism

The imperative view holds that *e*'s painfulness is reduced to the fact that *S* receives a command from her relevant sensory system to the effect that *S* engage (or stop engaging) in a certain sort of behavior directed to the bodily location where the pain is felt. Imperativists differ on whether pains are entirely constituted by commands (Klein) or also contain sensory information (Hall and Martínez). For our purposes, nothing much hangs on this issue, since on *both* accounts it is supposed to be the commands that explain the motivational and rationalizing role of pains. The sensory content is at best value-neutral. So on all imperative accounts it is the *command* that is supposed to give a reason for our behavior.

We have been assuming that it makes sense to talk about experiential representations with the general understanding that the phenomenology of perceptual/affective experiences can be reductively explained by appeal to such representations and the way they are processed. Imperativism appeals to experiential *commands*—experiential representations in imperative mood. We will assume that this view is different from views that appeal to experiential *desires* or causal/functional roles—e.g., phen-desires—such as ours where it is the psychofunctional role (m-processing $\approx$ CR, above) of incoming sensory information that turns the resulting mental process into a desire-like process directed toward the object of the experience about which the information is gathered through sensory means.

The difference is, roughly, analogous to the difference between the following two scenarios.

The police shout at the protesters: "Disperse and leave the square immediately!" This is a serious command issued by a (usually) legitimate authority. This command may or may not be obeyed by the protesters. But whether it is obeyed is not a direct causal consequence of the command. The command is just one element in the network of causes that result in the actual dispersion of the crowd—if they decide to disperse.

But now suppose the protesters don't comply, perhaps because they fundamentally disagree with the rational ground of the commands or they want to make their voices be heard more strongly for their cause, or whatever. So the police start using powerful water cannons on the protesters and fire tear gas at them. Despite disagreeing with the rational grounds for the police action, the protesters now have a perfectly good reason to leave the square! In fact, in a certain sense, they have no choice in the matter: we can imagine that the water cannons are so powerful that they are almost sufficient to physically disperse the crowd.

Clearly, the police action (shooting high-velocity water streams in big volumes at the protesters) is no longer a command (except perhaps *metaphorically*). Rather, it is an act to *force* the crowd to disperse. As such, once executed, it may fail or succeed in *causing* the crowd to disperse. The police action has an immediate causal/functional efficacy that the command lacks.

We consider the imperativist view of issuing commands in sensory code to be *analogous* to the police command and the experiential desires to the police action. The latter is essentially a matter of causal/functional role. The former is not.⁴⁰ When you look at the candidate description of CR we tentatively gave above to illustrate the kind of psychofunctional role the incoming sensory information must be subjected to in order to constitute affect, you will see that m-processing, although it may not often result in actual action, is nevertheless inherently causally efficacious in affecting and biasing all sorts of learning, motor, and decision-theoretic mechanisms—it *changes* processes in these mechanisms and the way they interface with behavior.

It is possible that the defenders of imperativism are open to interpreting their view as compatible or indeed requiring some form of psychofunctionalism. Indeed, Hall (2008) as quoted above seems quite friendly to the idea of cashing out the notion of a command functionally. Similarly, Klein writes: "An animal who has not had food or drink in days does not need to be informed about her body. She needs to be driven to eat and drink; a biologically effective system for doing so should not bother with the niceties of explaining why. So, too, with pain" (2007: 526). This notion of being "driven" and of coercing action without an intermediary explanation seems thoroughly functionalist rather than representationalist. We think that the *motivational oomph*⁴¹ thus understood cannot be captured by the notion of a mere command to the experiencer. However, if the notion of command assumed by imperativism reduces to a certain kind of causal-functional role, then it is at best misleading to advertise the view as a form of (strong) representationalism—it is at best a form of weak intentionalism where the notion

of sensory information plays an important role in the ultimately functionalist explanation of sensory affect, and the talk about commands or imperatives are at best metaphorical—just useful aids for understanding the psychofunctional role the sensory information plays to count as affective. The all-important notion of intentional *content* does not do much, then—the motivational punch is delivered by the functional role, and the view collapses to a form of (psycho)functionalism. In our evaluation, we will thus assume a version of imperativism that is not functionalist but thoroughgoing representationalist.⁴²

With these initial clarifications in mind, let's have a look at the explanations offered by imperativists. With the appropriate substitutions, two of the four claims that need explaining become:

- (d-Motivation_{imp}) *e*'s incorporating the command <see to it that *d* is no more!> to Sally *inherently* (but defeasibly) motivates Sally's *d*-directed behavior, which is to say that it defeasibly motivates it all by itself without the need of any beliefs and desires.
- (d-Justification_{imp}) *e*'s incorporating the command <see to it that *d* is no more!> to Sally, insofar as *d* is in fact bad for her, *provides* Sally with a good reason for her *d*-directed behavior—it *justifies* her *d*-directed behavior (as well as her belief that *d* is bad for her, d-Belief_n).

Let's start with a few observations. First, note that the commands almost never incorporate the *rationale* for their own issuance. The rational compliance is almost always a matter of an existing setup or understanding of background conditions and knowledge of the social, cultural, and physical context. Second and consequently, the reception of ordinary commands rarely (if ever) motivates and rationalizes actions all by itself. A response to a command by whoever receives it is almost always a complex function of the receiver's background beliefs about the authority, legitimacy, or reliability of the issuer of the command combined with the receiver's background concerns, her perception of the context, as well as her standing and occurrent desires and preferences (and her mood). It is this set of background beliefs and concerns that determines whether the reception of command will succeed *motivating* as well as *justifying* the receiver's actions (and the subsequent changes in her beliefs and desires). If this is true and applicable to commands issued in sensory code, (d-Motivation_{imp}) and (d-Justification_{imp}) are just false—imperativism cannot explain (d-Motivation_{imp}) and (d-Justification_{imp}).⁴³

Indeed, remember the propositional attitudes Sally forms upon burning her finger:

- (d-Belief_i) the belief that she burned her finger (i.e., *d* occurred), and
- (d-Belief_n) that this is bad for her (i.e., *d* is bad for her), and
- (d-Desire) the desire to get rid of *d* (i.e., that *d* cease or be no more).

In evaluating evaluativism, we have assumed that *e*'s unpleasantness is both the etiological and justificatory ground of d-Belief_n. Leaving aside how "bad" is to be

understood, this is an intuitively plausible assumption: the experience seems to present *d* as being bad. The intuitive plausibility of this assumption under imperativist interpretation becomes moot, however: it is quite unclear what the reception of a mere command to the effect that *d* cease has anything to do with Sally's belief that *d* is bad for her, especially in the second imagined scenario when Sally forms a belief about *g*—the gentle touch when she is suffering from allodynia:

(B) *this* is bad for me (i.e., *g* is bad for me).

We take it that B is true, when believed by her in the imagined context, whether or not she knows *g* is physically harmless. Again, it is not clear how a command in sensory code to the effect that *g* stop could rationalize and justify B—unless combined with a lot of background beliefs and concerns and a heavy reliance on inference, but this would falsify (d-Motivation) and (d-Justification).⁴⁴

Similarly, the rationality and warrantedness of Sally's original desire (d-Desire) as well as her second desire,

(d-Desire*) the desire to get rid of *g* (i.e., that *g* cease or be no more)

become at best quite *indirect* and *inferential* depending on what else Sally knows about the physical context, the reliability of the relevant commands, and what else she believes and prefers.

But most importantly, on imperativism, it becomes quite mysterious why Sally forms:

(e-Desire) the desire that the pain experience, *e*, itself stop, and
(e-Belief_n) the belief that *e* is bad for her.

We take it that these are *directly* justified and warranted by Sally's experience, *e*. And so are the following by her experience, *e**, in the allodynia case:

(e-Desire*) the desire that the pain experience, *e**, itself stop, and
(e-Belief_n*) the belief that *e** is bad for her.

But we are in the dark about how imperativism can imply or accommodate these truths.

With respect to Sally's *behavior*, things are worse. To be a *good* reason, the content of the command has to align with the goal of our actions. This means that the command has to reference or "make intelligible" our actions *with respect to d*. As we have been assuming, Martínez's version of the form of the relevant commands does just this: <see to it that *d* is no more!>

Is this a *good reason* for Sally? It does not appear to be, because it is entirely grounded by *d* itself (and/or Sally's epistemic access to it). That is, Sally *already* has a good reason to remove the disturbance: she has a disturbance in her finger, and she's aware of it! What more reason does the addition of the command add?

Well, none. This is clear from pathological cases: hold fixed the command but remove the disturbance. Does Sally still have a good reason to behave so as to remove *d*? Clearly not. Reverse the case: keep *d* but remove the command. Does Sally have a good reason to remove *d*? Yes. We thus have a double-dissociation, and it cannot be the command that supplies the reason for Sally's *d*-directed behavior. The situation is even clearer in Sally's *g*-directed behavior. Note that such dissociations are not rare and hold no matter what semantic account of imperatives is used to cash out the view.

This, of course, just brings us to what imperativism has to say about the remaining two explananda under the appropriate substitutions:

- (**e-Motivation**_{imp}) *e*'s incorporating the command <see to it that *d* is no more!> to Sally *inherently* (but defeasibly) motivates Sally's *e*-directed behavior, which is to say that it defeasibly motivates it all by itself without the need of any beliefs and desires.
- (**e-Justification**_{imp}) *e*'s incorporating the command <see to it that *d* is no more!> to Sally is *intrinsically* bad, which is to say that it *constitutes* a good non-instrumental (i.e., justifying) reason for both Sally's *e*-directed behavior and her anti-*e* desire (e-Desire) and belief (e-Belief_n).

Clearly, the imperativists have even more difficulties with these. To provide a good reason, the content of the command again has to line up with the command. But how is this supposed to work? Must the command to Sally be something like: <see to it that *e* is no more!>? But what is this command she experiences? Well, it's nothing more than *e* itself! So *e* is an experiential command to see to it that that very command is no more. How can such a self-defeating command have any reasons-giving force? It is *like* an officer ordering a cadet: Disobey this command!⁴⁵

What can the imperativist say here? Martínez (2013 SSPP) suggests that "it's not like pain is whispering its command in my ear, and I get to judge its authority, or its urgency. Nothing prevents pain from simply changing its subject's TDL from the inside, so to say." The TDL, on his view, is a to-do list of all the items one is supposed to do.⁴⁶ A command is then cashed out as something that adds an item to the TDL. But why should adding an item to my TDL give me a reason for anything? Suppose Sam assiduously keeps a to-do list on his desk. When he's not looking, a grad student adds <Give Jones an A!> to his list. Does its presence there give Sam a reason to give Jones an A? It certainly doesn't seem so. Martínez suggests a possibility: "The reason [why it is rational to try to stop the pain when there is no physical threat] is clear: the pain imperative is clogging our TDL, forcing us to attend urgently to seeing to it that a certain bodily disturbance is no more. If we know on other grounds that we do *not* have to see to it, it's entirely rational and justified to try to prevent the pain from messing with our to-do list."

The nature of Martínez's suggestion is analogous to the difference between the police action and the police command we introduced above. Commands

themselves have limited reasons-giving force unless wedded to some downstream story (clogging our inbox). If the nociceptive information processing is directly connected to the ordering of one's TDL in a way that cannot be more or less rationally/cognitively controlled, it is no longer a command. To go this route is to abandon imperativism, and with it, representationalism. But even worse, this suggestion comes down to admitting that a psychofunctionalist account is the right explanatory account for the unpleasantness of pains. In this story, it is not the *commands* themselves that undergird the motivational oomph, but rather the hard-wired causal role of incoming sensory information in *forcing* the priorities to be ordered differently—this is psychofunctionalism. Furthermore, it is not clear to us how literally we should take this talk of “to-do list.” When taken literally, this seems to require that subjects keep a to-do list for themselves keeping track of which commands should be issued to oneself in which order. But why should one have a list of self-commands? All the issues that we raise about commands in ordinary contexts arise here too. What reasons might one have to obey such commands one issues to oneself, etc.? It appears that the talk about TDL is just loose talk about having beliefs, desires, preferences, intentions, decisions, volitions, etc., or whatever the appropriate subpersonal computational decision-theoretic mechanisms and processes are involved.

Finally, Martínez's suggestion comes down to a rejection of (e-Justification). In effect, Martínez tells us that pain experiences aren't *intrinsically bad*, but they are bad for their inconvenient consequences—they, for instance, tend to clog our to-do list. (If, on the other hand, Martínez suggests that this is not a *consequence* of our painful experiences but rather is part of what it is to have a painful experience, then again, the suggestion is not imperativist anymore, but psychofunctionalist—and we are sympathetic.)

5 Conclusion

Despite our criticisms, we believe that all three theories of sensory affect present views that are close to each other. All are naturalistic and offer a *posteriori* accounts of sensory affect playing with more or less the same naturalistic resources available to them. We believe that at the core of these resources is the functional/causal role that the incoming sensory information plays. We labeled this complex set of processes defined over sensory information *m-processing* (with a role like CR), and identified it, roughly, with a more or less personal-level experiential desire, *phen-desire*, directed to the physical stimulus responsible for the relevant sensation. The satisfaction or frustration of this *phen-desire* as registered by the sensory component is meant to capture the positive or negative affective phenomenology of sensory experiences.

We are not anti-representationalists. Although we reject strong representationalism à la Dretske and Tye, we welcome information-theoretic treatment of sensory and cognitive processes. We think that sensations represent worldly properties and magnitudes, but their phenomenology is not exhausted by their representational content. This is particularly true for affective phenomenology,

which, according to us, is not representational at all. The metaphysics of sensory affect is the functional role the sensory information plays in the mental economy of the agent. Of course, experiential desire or aversion towards a stimulus needs to represent the stimulus. This is done by the sensory signal whose personal-level registration is the sensation (phen-belief), but the mechanism responsible for adding an affective dimension to the sensation is the m-processing of that very *same* signal—its causal/functional role, CR. It is not clear to us that there is any serious alternative to this basic framework. Indeed, it is not clear to us that imperativists or evaluativists do or can reject this framework.

Imperativists seem closer to our view in many ways, but they muddle the core issues by talking about the imperative content or commands, and then worry about developing a semantics for these as if the basic philosophical and metaphysical puzzles surrounding sensory affect would be solved if we had a working semantics for imperatives or commands. This may be a useful thing to have, but we don't see how this project can succeed without essential appeal to the functional/casual role of sensory information to make this semantic story work as a metaphysical account of sensory affect as instantiated in functioning agents.

Evaluativists, on the other hand, need to tell us three things. First, what is it for a physical disturbance to be bad in the way that is required for its experiential representation as bad to be accurate or inaccurate? If we are dealing with a notion of badness other than relational badness (r-badness) or intrinsic badness, we do not have any clue about what it might be.⁴⁷ We have been pretending to know what physical/objective/bodily badness comes to, but our criticism above shows that we do not—neither do evaluativists, it seems to us... Secondly, they need to tell us how representing a physical condition as bad, in whatever physical/bodily sense they have in mind, can motivate in the absence of any beliefs and desires (d-Motivation) without relying on motivational internalism. Motivational internalism, in the context of *sensory affect*, is a mysterious thesis: how could experientially representing something as physically bad all by itself have an internal (necessary) connection to motivation? Unless we assume a functionalist framework of the sort we defend, the existence of such an internal connection is mysterious and puzzling. But if, on the other hand, we assume our functionalist framework, evaluativism becomes otiose.⁴⁸ Thirdly, and perhaps, more importantly, how is representing some bodily disturbance as bad itself bad? This is the notorious problem of messenger-shooting for all representationalists. The problem is particularly difficult if evaluativists have some sort of objective physical/bodily badness in mind.

This problem, in many ways, is at the very core of philosophical puzzles about sensory affect—it is not just a by-product of certain ways of trying to give a naturalistic account for sensory affect. We are puzzled that all representationalists, while trying to bring philosophical illumination in all sorts of ways, seem to leave the main difficulty more or less untouched.⁴⁹

Notes

- 1 For distinctive feeling views see (among others) Moore (1903/1993: 12), Brink (1989: 221), and Bramble (2011). For the hedonic tone view, Broad (1930), Duncker (1941), Kagan (1992), Sprigge (2000), Crisp (2006), Smuts (2011), and Labukt (2012), among others.
- 2 Defenders of attitudinal views include Alston (1968), Davis (1981, 1982), Brandt (1979), Feldman (1997, 2004), and Heathwood (2007), among others. Parfit (2011) and Brady (2013) develop the attitudinal theory in terms of liking (or disliking) a sensation, where this hedonic attitude is taken to be different from desiring in that it is supposed to be not responsive to reasons. It is difficult to classify Sidgwick (1907): he offers what appears to be an attitudinal theory, but he uses “desirable” rather than “desired” (1907/1981: 127).
- 3 See, for instance, Armstrong (1962, 1968), Pitcher (1970), Tye (1995), Hall (1989), and Kahane (2009). Robinson (2006) gives an account of sensory pleasure that is hard to classify but seems closer to the attitudinal theories, as he takes the pleasantness to essentially involve intentional and evaluative directedness toward sensations.
- 4 Even this much may not be agreeable to everyone. There are theorists who think that experiential affect is the perception of objective value out there. Perhaps they think that its phenomenology is as robust—whatever phenomenological robustness comes to. It’s hard to tell. See Johnston (2001).
- 5 Defenders of representationalism include (among many others) Tye (1995), Dretske (1995), and Byrne and Tye (2006).
- 6 See, however, Tye (2005) and Aydede (2005, 2009a).
- 7 Klein (2007) uses “intentional” rather than “representational,” reserving the latter for only descriptive/indicative content. We will continue to use “representational” as having (conceptual or non-conceptual) propositional content in whatever attitude or mood the representations are deployed, and use “representationalism” to label the view that all experiences are representational.
- 8 By invoking Millikan’s consumer semantics as a natural avenue for naturalizing imperative content, Martínez also seems to implicitly support the reduction to psychofunctional role—see below.
- 9 Versions of this view have been held by Bain (2013), O’Sullivan and Shroer (2012), Cutter and Tye (2011, 2014), and Tye (2005). An early and influential version of this view seems to be Helm’s (2002) *Felt Evaluation* view—although Helm seems to develop and defend his view on *a priori* grounds.
- 10 The psychofunctionalist, unlike the analytical functionalist, need not specify a theoretical identity or a set of necessary and sufficient conditions on the explanandum right at the beginning. Instead, the theorist can remain neutral about the ultimate character of the theory and simply proceed with functional analysis. The task is one of reductively explaining a complex mental state by breaking it down into its functional parts, and then showing how those various parts work together to produce the higher-level state (see Cummins 1981). We can decide on the nature of the identities and linking propositions after the spade work has been done, as a kind of mopping up operation. Importantly, while there need not be theoretical identities at the lowest levels (for mundane reasons having to do with multiple realizability and functional constancy), at the higher levels of analysis, where the philosopher works, there will be something like a reductive *a posteriori* identity between the target state and a (higher-level) functional role. It is precisely at this level that we plan to couch our theory. We do this by coordinating person-level and empirical work to define a set of high-level roles constitutive (or at least, characteristic) of sensory affective states—see Figure 2.1. See also Clark (2005), who is a fellow psychofunctionalist, about the painfulness of pains. Tim Schroeder’s (2004) work on pleasure is also generally in

the same psychofunctionalist ballpark—although he seems to think of himself as a representationalist.

- 11 See Berridge (2004), Rolls (2005), and Panksepp (1998). We do not take these bullet points to exhaust CR, nor do we take them, individually or even jointly, to be very illuminating. The point is to show intuitively the sort of psychofunctional causal role we have in mind. CR is for now a placeholder for whatever the ultimate scientific account of this role will turn out to be. In future work, we hope to fill in the system-level empirical details.
- 12 There is growing evidence that affective processing of sensory information is done in both parallel and serial fashion in mutually interacting distinct computational modules. For instance, the mechanisms responsible for computing the reward value of a stimulus, for the utilization of this information for learning, and for the execution of behavior subsequent to this, seem to be distinct, and there is strong evidence that they can and do come apart. See, for instance, Pecina *et al.* (2006).
- 13 Brandt (1979: 38) gives this as a definition of sensory pleasure.
- 14 Compare to Byrne's *exing that P* (2009)—roughly *experiential/phenomenal believing that P* that feeds into the mechanisms of ordinary believing under certain conditions. See also Aydede (2000) for a similar explanation of m-processing as, what he calls, “desiring*.”
- 15 See Aydede (2014; forthcoming) who likes to explain the phenomenology of this affective modification of incoming sensory information in adverbialist terms.
- 16 See Aydede and Fulkerson (2014) for the distinction between the Object and Experience Views that locate the affective qualities differently. In their (2013), they defend a Lockean dispositional account of affective qualities as attaching to the worldly objects of our experiences.
- 17 Note that none of these theories are meant to solve (all by themselves anyway) the general problem of the explanatory gap said to exist between phenomenal consciousness and the physical processes that seem to support/realize this consciousness. So we will leave this issue aside—see Aydede & Güzeldere (2005) for a detailed proposal about how to address this general problem.

Another potential worry is that psychofunctionalism may not be in competition with the other representationalist views, since, one might argue, any naturalistic representationalism is committed to some psychofunctionalist realization or other—it all depends on what actual psychofunctional role (more complete and precise specification of CR) is being proposed as doing the explanatory work. We set this worry aside too. This is partly because we strongly suspect that any further elaboration of this functional role along the way CR suggests above won't be suitable as a reduction base for the representationalist proposals expressed by the right-hand side of these identities. Another reason is that we suspect that any proposal about how to naturalize representational content will have to appeal to more than (narrow) functional/causal roles. But this is a huge topic we cannot take up here.

- 18 Depending on the context, *d*-directed behavior can be protective, rehabilitating, comforting, tending, terminating, stopping, etc. We will sometimes use “*d*-avoidance behavior” for this constellation of behavior patterns.
- 19 These are ordinary beliefs and desires with conceptual content. For convenience, we use “*d*” and “*e*” constructions to indicate that these beliefs and desires are directed towards the disturbance, *d*, and the experience, *e*, respectively. We use subscripts “*i*” and “*n*” as gnomic devices to mark that the belief is indicative (descriptive) or normative (evaluative), respectively.
- 20 And the belief that this is *why* she is feeling a *pain in her finger*. This sort of belief attributing pains to body parts is particularly problematic for representationalists. But we will leave this issue aside in this paper (see Aydede 2009a).
- 21 Cf. Bain (2013) and O'Sullivan and Shroer (2012). In explaining intentional behavior, we follow the Davidsonian tradition in assuming that causal and rationalizing

explanations don't compete—they are different descriptions of one and the same process (*perhaps* at different levels).

- 22 That I believe my watch is in the drawer is a *motivating/rationalizing reason* for why I open the drawer. But given that my watch is in fact on the kitchen counter, it doesn't provide me with a *good reason* for my action: it fails to normatively *justify* my action by dint of being false. Also, in what follows, when we use "motivation," we will include *rationalization* in its intension.
- 23 That is, without the need for any beliefs and desires *additional* to the stock of Sally's beliefs and desires she already had roughly at the time at which she started having the pain. We are not claiming that it's possible to have motivating pains without possessing *any* beliefs and desires—perhaps only intentional systems can be properly *motivated*.
- 24 This doesn't rule out the possibility of painfulness being instrumentally good, and therefore desired.
- 25 Corns (2013) also seems to deny d-Motivation. But this is because she thinks, following the experimental results of Berridge and his colleagues (see, e.g., Pecifia *et al.*, 2006), that hedonic value (pure pleasantness) can be separated from the motivating role of certain sensations. However, the personal-level or phenomenological significance of these results is moot. Berridge and his colleagues certainly seem to have unearthed separate subpersonal mechanisms that underlie motivation, reward, and learning. But we have doubts about the phenomenological claims made directly on the basis of these mechanisms. See, however, footnote 35 below. Martínez (forthcoming), after distinguishing motivation from justification as well as damage-directed behavior from experience-directed behavior in a way similar to our four claims above, also seems to deny the last three claims—although he pitches his account explicitly in terms of whether *pains* can provide the relevant range of *reasons*.
- 26 Ironically, this example is taken from Bain (2013), where Bain criticizes imperativism as incapable of explaining why it would be irrational for Sally to obey the command commanding Sally to act against gentle touching.
- 27 In this case, though, the justification involved would not be direct and the reason provided by the experience's evaluative content would be instrumental.
- 28 Can the evaluativist claim that the evaluative content of painful pain experiences is that *d* is r-bad for Sally? In other words, the unpleasantness of Sally's *e* is just its experientially representing *d* as r-bad. But this would be circular. It just comes to the claim that the unpleasantness of *e* is just representing *d* as causing an unpleasant experience, *e*. (Cf. Bain's own criticism of what he calls the Dislike View in his 2013.)
- 29 Again, if an evaluativist holds a relevantly modified version of motivational internalism, *perhaps* (e-Motivation_{eval}) may be retained without too much trouble. But motivational internalism is controversial, and we are not sure that evaluativists would be happy to rest the plausibility of their case on its truth.
- 30 This is, of course, *not* to deny that the physical disturbances are important and usually objectively bad for the organisms who register them.
- 31 In fact, it does not take us anywhere. See Aydede and Fulkerson (2014) for elaboration of this criticism of representationalism about affect.
- 32 See also Jacobson (2013) for a nice elaboration of this criticism. Cf. Aydede (2005; 2009b). Cutter and Tye (2014) respond to Jacobson by denying (e-Motivation) and (e-Justification).
- 33 Is it plausible to reject (e-Justification)—or (e-Motivation), for that matter—on the ground that the badness of pain experiences is completely extrinsic—for instance, by saying that it derives it from the badness of its consequences, such as being exhausting or distracting? (Cf. Cutter & Tye, 2014; Martínez, forthcoming). We cannot argue against its plausibility in this paper, but if evaluativism can be saved only by denying (e-Justification), it's not worth saving, especially when there are clear and better alternatives to it, such as psychofunctionalism. In general, we believe that our own

naturalistic view will be preferable to any other naturalistic view that denies that painful pains are intrinsically bad. We invite the reader to do a comparative assessment and leave the matter at that.

- 34 In a forthcoming article, in responding to an anonymous reviewer, David Bain seems to suggest that his evaluativist view can be formulated also as an experiential desire view where the desire is directed towards bodily disturbance, *d*. We note here that these views are not equivalent if the represented badness of *d* is physical or objective badness not constituted entirely by *r*-badness. If, on the other hand, the represented badness is *r*-badness, then evaluativism is circular, but there is no corresponding difficulty of an experiential desire view—indeed, it is not at all clear what an experiential desire has to do, *content-wise*, with the *physical* badness of a disturbance like *d* or *g*. Also, and relatedly, if these views were equivalent, there would be no ground for worrying about providing a psychosemantics for the evaluativist content. The experiential desire view, on the other hand, does not need a psychosemantics beyond the psychosemantics for the indicative content that *this* disturbance is occurring: being aversive to this content is a *functional* affair that does not need a psychosemantic explanation. Similarly, there would be no ground for worrying about the messenger-shooting problem: the experiential desire view delivers a quite natural and plausible account of why the experience itself is bad in a way evaluativism does not. But Bain, quite rightly, worries about these two problems deeply. (In personal correspondence, Bain has kindly pointed out that he is officially neutral about whether desires can be analyzed as evaluative beliefs of a certain sort and that our phen-desire view is indeed not equivalent to his evaluative content account of pain's affect.)
- 35 The qualification “very roughly” is required because the description in terms of phen-desires and phen-beliefs are close to personal level descriptions and may not neatly correspond to any natural delineation of mechanisms constituting *m*-processing. For instance, what is phen-believed may be that a certain taste objectively understood is present in my mouth as a result of, say, biting a ripe strawberry (descriptive sensory phenomenology). But this content is *also* what is phen-desired (affective/conative phenomenology). Are there further mechanisms during *m*-processing that additionally register the fact that what is phen-believed satisfies what is thus phen-desired—a state analogous to a meta-belief? Whether this be so depends on further empirical facts about how the affective system works as a network between informational, motivational (drive), and learning mechanisms. We won't further explore the details here. Suffice it to say that the phen-belief and the phen-desire involved in a given episode of sensory affect are functionally defined over the same sensory signal carrying information about, more or less, what is happening at the location of the relevant sensory receptors. Without a common binding and the availability of this information, it is hard to fathom what the point of having an affective system may come to. We do suspect, however, that the consciously registered (dis)pleasure may arise out of an engineering need to have an explicit signal for desire (dis)satisfaction: an autonomous system capable of individual reinforcement learning needs to know when and how its desires or phen-desires are satisfied or frustrated as a consequence of its own actions.
- 36 We are thus in agreement with Goldstein (1980, 1987) that the intrinsic badness of pains and the intrinsic goodness of pleasures are manifest to rational animals (who, we would like to add, have the requisite conceptual resources).
- 37 A further substantial advantage of this view is that it may offer a plausible account of sadism and other cases of disconnect between pain and our all-things-considered desires.
- 38 Additionally, our psychofunctionalism seems to provide a potential resolution of certain core aspects of the debate between (normative) reason internalism and reason externalism. It is hard to say without generating controversy what these views are, and we will not attempt it here. Suffice it to say that the latter view, as we understand it, implies that all desires admit of independent justification, while the former denies

this. So, suppose we have naturalistically secured the notion of experiential desire and belief—as indicative and conative/desiderative representation in sensory code (as sensory information being m-processed). Then just as sensory/perceptual experiences can justify our perceptual beliefs, experiential desires (phen-desires) can rationalize and justify our desires as conceptually structured propositional attitudes (call these, c-desires). Reason externalism could be defended only as applied to c-desires—all c-desires do or can admit non-arbitrary independent reasons. For instance, the reason for c-desiring to get rid of pains is that pains are intrinsically bad (continually frustrated phen-desires) and we can grasp them as such (cf. Goldstein 1980). But then there wouldn't be any reason internalism as applied to phen-desires—token phen-desires just aren't the kind of states that can have reasons unless they are perceptions of objective sensuous value (à la Johnston, 2001), in which case they would become phen-beliefs about objective sensuous values—but we set this view aside here. Reason externalism would thus win the debate, but the insights of reason internalists would nevertheless be preserved—again, assuming that sensory desiderative phenomenology is not the registration of objective value beyond one's skin. Importantly, we can also objectively explain the existence of phen-desires themselves (their *raison d'être*, in other words) on the basis of how or why evolutionary selective pressures worked out the way they did. This may amount to their “justification” as types—but this sort of justification is not at issue in the debate between these two parties.

- 39 Furthermore, the intrinsic value we have in mind comes in degrees and doesn't compete with the sensory affect's also having instrumental or extrinsic value. In fact, for just the dialectical purposes of this essay, we may even restrict the intrinsic value to just non-derivative or final value *for the agent* whose experience has the affect.
- 40 However, the imperativist ought to give an account of the imperative mood: what makes a certain representation a command rather than an indicative? We think this can't be done without adverting to the functional/conceptual role (use) of the representational vehicle. It may be that once this role is spelled out fully by the imperativist, the view will collapse into a functionalism about affect. But here we leave this issue aside.
- 41 Thanks to Jennifer Corns (2013) for this wonderfully useful and insightful phrase!
- 42 Sometimes, imperativists seem to present their view as offering a different notion of content, *imperative content*, rather than a command. The point of the emphasis seems to be that the explanatory work is done solely in terms of content—it is just that there is more than one notion of content: indicative, imperative, interrogative, etc. These are all contents. Compare this to saying that there is one kind of content (propositions) but that this content can be put in different uses by the different ways their representational vehicles function. There are larger issues about which way is a better way of proceeding from a theoretical/explanatory point of view. Clearly we think the latter way is a better way and find the former somewhat mysterious. But we cannot discuss these larger issues here.
- 43 Klein (forthcoming) seems to reject (d-Motivation), and thus, (d-Motivation_{imp}). Following Helm (2002), Klein agrees that commands all by themselves don't motivate in the absence of the subject's care and concern about the immediate physical integrity of her body. But such care and concern certainly entail that the subject has a *desire* to protect her body. It seems likely that Klein thinks that this desire is hard-wired or is a standing or dispositional desire of some sort. The trouble with this suggestion is that if the subject has a desire of this sort, all she needs is the information that her body is being physically threatened, and this information is already (experientially) supplied by the sensory aspect of the pain. Commands are not needed. (Klein doesn't think that pain experiences have any non-imperative content. But this makes things more difficult for him, not less. So we are putting this option aside.) The proposal then at best comes down to the experiential equivalent of a belief-desire explanation/rationalization of the subject's subsequent *d*-directed behavior—there is no mystery

about how such states can motivate and rationalize, and there is no explanatory need for commands or imperative content.

- 44 On Klein's austere version of imperativism, *e* doesn't incorporate any direct sensory-discriminative (indicative) representational content about the condition of the body. So on this view, a belief like (d-Belief_i)—and probably the desire (d-Desire)—is not *directly* justified by the content of the experience. Experience (i.e., the command in sensory code) plays only an indirect justificatory role when combined with background assumptions and preferences in the way in which a soldier can come to inferentially know that the hostilities started upon receiving his marching orders (see Klein 2007: 523). Furthermore, on Klein's view, commands aren't really about the (actually or potentially) damaged parts of the body, but rather they are about *how not to use* the body parts involving the damage. This further distances the pain experience from the damage or the nociceptive physical stimulus. On Klein's view, pain experiences have very little to do with what is harming or seems to be harming the tissue where the pain is felt. It appears that the sufferer's knowledge of where the pain is felt is also inferential depending on what body parts are commanded not to be used in a certain way and the background knowledge of the sufferer. We find the austere version of imperativism to be implausible, and for that reason, have been assuming a version of imperativism (along with Hall and Martínez) that does not deny the sensory-discriminatory content of affective sensory experiences. Indeed, we are not sure it even makes sense to be an austere imperativist about positive affect involved in various gustatory, olfactory, tactile *pleasant sensations*.
- 45 Or, as pointed out by David Bain in correspondence, more accurately like <stop this command!>. Imperativists, however, have never invoked experience-directed commands.
- 46 Martínez, in his paper (forthcoming), has dropped the analogy with “to-do-list,” but the basic idea behind it is still in force. In his forthcoming paper, which came to our attention considerably after we'd delivered the present criticism at the 2014 Pacific APA, Martínez explicitly addresses the question of how pains can give or be reasons for behavior. But, as far as we can see, our main criticisms above remain. Furthermore, Martínez seems to deny (d-Justification), (e-Motivation), and (e-Justification). We say “seems” because he is relatively clear only about denying (e-Justification). But we fail to see how he escapes denying the other two, given what he says about what the painfulness of pains comes to and how it motivates or justifies.
- 47 See our (2014) for a criticism of a proposal by Cutter and Tye (2011). Our criticism extends to any proposal that identifies this badness with a completely objective (extramental) property—however complex (historical, relational, etc.) it may be.
- 48 Bain (forthcoming) claims that general care about one's body is an existence condition for one to have painful pain experiences, i.e., for one to have pain experiences that represent disturbed bodily conditions as bad. We find this stipulation also mysterious. If one cares about one's body, in whatever sense relevant here, then one has *desires* about not to have physical disturbances in one's body. But if so, we have a fairly standard (and more economical) explanation of the motivational power of painful pain experiences without relying on identifying the painfulness with experiential representation of disturbance as bad. We can just say that the painfulness is the frustrated experiential (or, otherwise) desires implied by this general body care. In fact, whatever the nature of the anti-damage desire implied by body care may be, when combined with the sensory information that bodily damage is occurring, we have a fairly straightforward explanation or rationalization of one's anti-damage behavior. It is not at all clear what we gain by insisting that the painfulness is the representation of disturbance as bad and that the body care and anti-damage desires are only existence conditions for such representations.
- 49 Cutter and Tye (2014) propose to “solve” the difficulty by denying there is any difficulty in the first place: they reject (e-Motivation) and (e-Justification) as false.

A precursor of this paper was written for an invited symposium at the Pacific APA meeting in San Diego (2014) titled “Painful Pains, Yummy Tastes, Stinky Smells: Sensory Affect.” We thank the other participants: David Bain, Colin Klein, Manolo Martínez, Amy Kind, and the audience members, for their very helpful comments and criticisms that have usefully shaped the subsequent revisions of this paper. We also thank David Bain, Michael Brady, and Jennifer Corns for their comments on an earlier version of this paper, as well as for the invitation to participate in a series of conferences on pain in Glasgow and to contribute to this volume. Many of the ideas here have been influenced and helped by the ongoing interactions made possible by these meetings and continuing conversation as a result.

3 A neuroscience perspective on pleasure and pain

*Dan-Mikael Ellingsen,
Morten L. Kringelbach, and Siri Leknes*

Introduction

Differential behavioral responses to stimuli that threaten or promote survival and procreation are a central feature of organisms, from vertebrates (Kringelbach and Phillips, 2014) to invertebrates (Perry and Barron, 2013), and even unicellular organisms (Jeon, 1973, Buonanno et al., 2013), pointing to mechanisms subserving self-care as evolutionarily ancient, perhaps universal, features across the species. Newborn babies are able to differentiate between pleasant and unpleasant facial expressions a few days after birth (Lewis, 2000, Farroni et al., 2007), and human prenatal infants show nocifensive responses about 28-30 weeks after conception (Craig et al., 1993, Lee et al., 2005). The time of emergence of conscious hedonic feelings, both during evolution and during fetal development, is still debated (Mashour and Alkire, 2013). Nevertheless, such early differentiation between harmful and beneficial stimuli is at the very least a precursor for hedonic valuation, the categorization of sensory events into having positive or negative value.

While much of human behavior is geared towards seeking pleasant experiences over both the short and long run, individuals will also work to avoid aversive or painful experiences. Hedonic valuation of sensation helps guide the choice of behaviors to engage in or avoid, thus rendering hedonic processing essential for survival. In order to be useful, however, hedonic feelings and corresponding behavioral responses need to be relatively flexible depending on the current needs of the organism. The hedonic value of sensory experiences is therefore closely linked to homeostatic processes. While the taste of chocolate can evoke intense feelings of pleasure, the sensation, although arising from the very same stimulus, can change its value and become less pleasurable after having eaten too much. Furthermore, the surrounding context, conveyed as other sensory information, is also a crucial determinant of hedonic value. The meaning, and consequently the pleasantness, of a gentle touch can be fundamentally different depending on whether the toucher appears as attractive and friendly, as opposed to unattractive (Gazzola et al., 2012) or threatening (Ellingsen et al., 2014). In order to make successful decisions about the relative hedonic value of a stimulus, the brain must be able to incorporate this internal and external information into the equation. Although hedonic value influences behavioral decisions, there is also flexibility between (dis)pleasure and behavior. Albeit pleasant, it is possible

to stop oneself from having another bite of chocolate, e.g. in order to stick to the diet, a competing motivation.

The conceptualization of hedonic or affective valence as a fundamental aspect of sensation has been central in psychological theories of emotion for more than a century (Wundt, 1897; 1980, Watson and Tellegen, 1985, Larsen and Diener, 1992, Barrett and Bliss-Moreau, 2009). Furthermore, there are extensive similarities between the brain networks and endocrinology involved in the processing of pleasure and displeasure, such as pain, which adds relevance to the question of whether the brain uses a “common currency” for hedonic value (Cabanac, 1992, Ramirez and Cabanac, 2003, Leknes and Tracey, 2008). Influential heuristic frameworks outline “hedonicity” as a general dimension, inherent of all sensations, ranging from distress to delight, with indifference in the middle (Young, 1959, Cabanac, 1979). Similarly, pleasure and pain are often considered opposites of an affective continuum. Many sensations appear unidimensionally nice or nasty, making it relatively effortless to verbalize how much one liked or disliked this particular sensation. Moreover, pain and pleasure are often opposing forces. Pleasant experiences like delicious food (Kringelbach, 2015), music (Vuust and Kringelbach, 2010, Bernatzky et al., 2011), beautiful visual images (Kenntner-Mabiala and Pauli, 2005), pleasant touch (Liljencrantz et al., 2012, Mancini et al., 2014), and perceived support from others can dampen concurrent pain. Conversely, the occurrence of pain often suppresses positive feelings, as illustrated by the burden of depression and anhedonia that frequently accompanies chronic pain conditions (Marbach and Lund, 1981, Marbach et al., 1983, Elvemo et al., 2015).

Nevertheless, there is still debate about how hedonic valence is coded in the brain, i.e. whether negative and positive hedonic valence are coded by i) separate neural systems that usually are negatively correlated but can co-activate; ii) within a single system where graded responses represent positivity or negativity on a “hedonic scale”; iii) or within a more flexible, integrated set of brain systems (Lindquist et al., 2015). Moreover, the existence of mixed emotions (Doyle et al., 1993, Larsen et al., 2001), which are associated with both pleasure and displeasure, such as tickles, bittersweet memories, or guilty pleasures, raises the question of whether it is meaningful to classify sensations along a unidimensional hedonic continuum (Soderpalm and Berridge, 2000, Larsen et al., 2003).

In this chapter we discuss evidence from animal and human studies of the brain circuitry that enables the brain to create pleasure and displeasure. We then review how top-down influences such as i) homeostatic utility, ii) internal motivational state, and iii) surrounding contextual factors interact with sensory signals to shape hedonic value and decisions. Finally, we discuss what mixed emotions can tell us about the nature of pleasure and displeasure in the brain.

1 The assessment of pleasure and pain

Measuring hedonic value in humans can consist of asking the person specifically about how much enjoyment or displeasure they experience, or about their overall affective tone. Subjective state ratings reflect affective processes that are available

to introspection, moderated by the motivational state of the participant (e.g. demand effects, malingering). Ratings most often occur along a single scale such as pain intensity or pleasantness-unpleasantness, thus restricting reports of complex and subtle hedonic experiences to just one or two dimensions. Subjective ratings are nevertheless central in the study of hedonic feelings, and indeed the subjective awareness of pleasure and pain are highly relevant. Without subjectively experienced pain, patients are less likely to seek medical care. Similarly, the absence of subjective enjoyment of rewards is an integral part of psychiatric conditions such as substance dependence and depression (Romer Thomsen et al., 2015). The challenges of obtaining reports of hedonic value in non-verbal humans and other animals have necessitated more objective measures of hedonic processing.

Functional neuroimaging studies that measure brain activity during self-reported enjoyment of rewards – be they sensory, social, or abstract rewards – have revealed many regions that play roles in pleasure processing, most notably the subcortical structures Nucleus Accumbens (NAc), ventral pallidum, and the amygdala; mid-brain structures Ventral Tegmental Area (VTA) and Periaqueductal Grey (PAG); and cortical structures orbitofrontal, cingulate, and insular cortices (Kuhn and Gallinat, 2012). However, since these neuroimaging techniques are correlative in nature, they cannot answer whether a brain region actually generates pleasure, or if it represents a readout of hedonic value in a subsequent process that primarily serves a different function. However, as we discuss later in this chapter, studies using microstimulation of these regions in rodents provide insight into the causal relationship between activation of specific regions and hedonic behavioral responses.

The subjective experience of pleasure is tightly linked to the concept of reward – indeed, the terms pleasure and reward are sometimes used interchangeably. However, hedonic processing can be conceptually divided into at least three components: wanting, liking, and learning (Berridge and Kringelbach, 2008, Smith et al., 2011). There is now substantial evidence that these three processes constitute partly separable neuroanatomical and neuropharmacological systems (Berridge and Robinson, 1998, Berridge and Kringelbach, 2013). Furthermore, hedonic reactions can be divided into subjective and objective aspects. While subjective experiences of pleasure can be indirectly assessed experimentally through self-report in humans, only objective reactions can be studied in animals. Wanting, liking, and learning aspects of rewards are associated with both subjective and objectively measureable “markers”.

Implicit – or core – “wanting” can be studied by observing how hard an individual will work to obtain a reward, while explicit wanting corresponds to the subjective experience of desire and cognitive goals (Berridge and Kringelbach, 2011). Measures of “wanting” include quantification of consumption of food, drink, or drug rewards (Comer et al., 1999), progressive ratio (PR) reinforcement experiments (e.g. Papaleo et al., 2007), and conditioned place preference (CPP) tests (Mayo et al., 2013). Correspondingly, motivational aspects of negative hedonics, like painful or disgusting events, can be studied by observing how much an individual will work to avoid a punishment (implicit), or by assessing subjective feelings of dread (explicit). In humans, the cold pressor test assesses willingness to endure pain

from maintaining one's hand in ice-cold water. Rodent models of pain avoidance behaviors include the tail flick test, and more recently, CPP measures of preference for pain relieving agents (Navratilova and Porreca, 2014).

While explicit liking or disliking corresponds to the subjective feelings of pleasure or displeasure, core “liking” or “disliking” can be assessed through observation of hedonic reactions during consumption of a reward or during a punishment. For instance, parents usually determine whether their pre-verbal toddler liked or disliked a given taste based on observation of their behavioral reactions. In human infants, facial expressions like licking the lips or making gapes are common reactions to sweet or bitter tastes, respectively (Steiner et al., 2001) (Figure 3.1b).

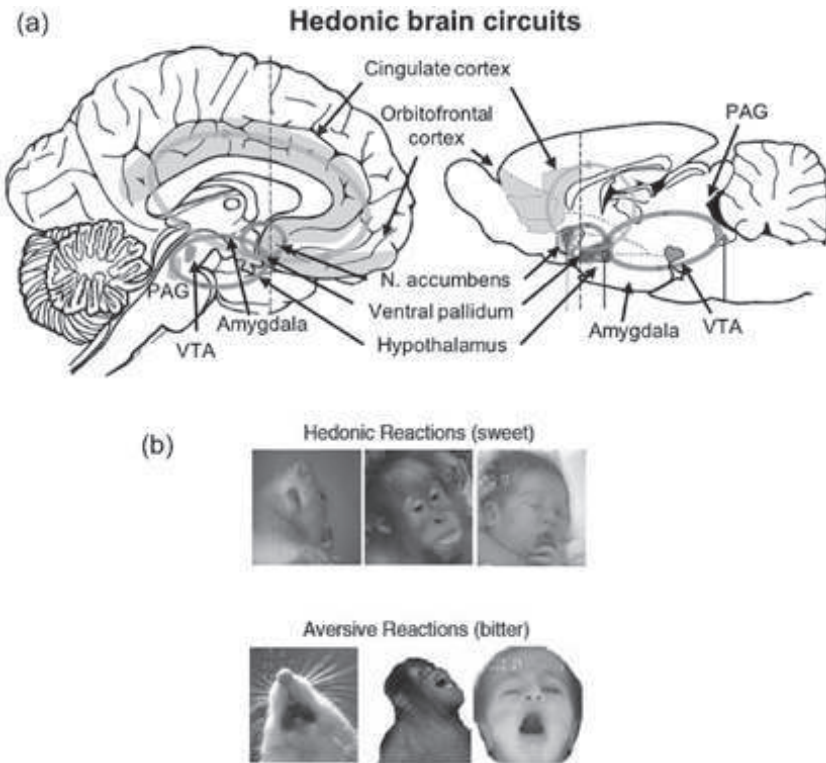


Figure 3.1 Brain circuitry for generating hedonic value. This schematic figure summarizes subcortical and cortical systems for generating and coding of hedonic value. (a) Brain circuitry involved in hedonic processing in rodents and humans. (b) Typical facial reactions to sweet and bitter taste are comparable in rodents, primates and human infants, and provide a useful model for investigating “liking” and “disliking”.

Source: (a) figure 2 (page 482) from Kringelbach, M. L. and K. C. Berridge (2009). “Towards a functional neuroanatomy of pleasure and happiness.” *Trends Cogn Sci* 13(11): 479–487. (b) figure 2 (page 11781), from Pecina, S. and K. C. Berridge (2005). “Hedonic hot spot in nucleus accumbens shell: where do mu-opioids cause increased hedonic impact of sweetness?” *J Neurosci* 25(50): 11777–11786.

This can be quantified by the intensity or frequency of licking or gaping responses and is also evident in both nonhuman primates and rodents (Steiner, 1973, Steiner et al., 2001). Such objectively observable hedonic responses are useful for conspecifics, as they provide cues to the other individuals about the hedonic value of the food reward. Similarly, pain is associated with automatic facial expressions that are likely to have evolved because of their capability of drawing attention from conspecifics who can potentially provide sympathy, support, and care (Williams, 2002), and learn from the experience of others. Recently, rodent facial expressions and body postures indicating pain are also receiving more attention as potential indications of “disliking” in rats and mice (Langford et al., 2006, Nakashima et al., 2015).

Using core “liking” responses as experimental outcome measures is useful, especially in animals and pre-verbal children where subjective experiences cannot be directly assessed. If a manipulation successfully changes an animal’s “liking” response to a reward, this indicates that the manipulation altered a process responsible for generating “liking”.

Rewards and punishments serve important roles as “teaching” signals for future behavior, promoting behavior leading to a beneficial outcome, while suppressing behavior leading to adverse outcomes (Skvortsova et al., 2014, Fields and Margolis, 2015). Core “learning” comprises a range of processes involved in associative conditioning and implicit knowledge, while explicit learning includes the conscious updating of cognitive predictions.

2 The brain networks generating hedonic value

Pain is a salient unpleasant sensation associated (by the experiencing individual) with actual or potential tissue damage (Merksey and Bogduk, 1994), and is often contrasted with sensory pleasure (Leknes and Tracey, 2008). Painful and pleasant sensations are usually rich and multifaceted events, involving e.g. sensory-discriminative aspects (Auvray et al., 2010), motor responses (Perini et al., 2013), motivational processes (Salamone and Correa, 2012, Richard et al., 2013, Schwartz et al., 2014, Volders et al., 2014), attention (Eccleston and Crombez, 1999, Crombez et al., 2013), and contextual unspecific sensations surrounding the hedonic stimulus (e.g. the sight of the flame, the smell of burnt skin, etc., when getting burned by a fire) (Eippert et al., 2007, Colloca and Finniss, 2012). In search of the neural substrate (or substrates) responsible for the actual hedonic experience of pain or pleasure, the sheer hurt or bliss, a major challenge has been, and continues to be, to disentangle these different components.

Human functional Magnetic Resonance Imaging (fMRI) studies of pain report activations of a range of brain regions in response to experimental pain stimuli, including the thalamus, primary and secondary somatosensory areas, insula, dorsal anterior Mid-Cingulate (aMCC), prefrontal cortex, amygdala, and brainstem structures (Tracey and Mantyh, 2007, Segerdahl et al., 2015). Some functional neuroimaging studies have reported a segregation of brain systems that process the affective or hedonic aspects of physical pain and those that process

the sensory-discriminatory aspects (Rainville et al., 1997, Kulkarni et al., 2005, Auvray et al., 2010). Importantly, while these structures all play important roles in pain processing, they are also involved in the processing of non-painful, and even non-affective, sensations, and may be responsible for more generalized functions related to salience and attention (Mouraux et al., 2011, Iannetti et al., 2013). Nevertheless, by using information in fine-grained spatio-temporal patterns of fMRI activations within this “pain-responsive” circuitry, recent endeavors have been able to accurately differentiate processing of selected painful versus non-painful stimuli (Liang et al., 2013, Wager et al., 2013). However, thus far, this has proven successful only within very restricted contexts, e.g. painful vs non-painful heat induced by differences in stimulus temperature, but not by cognitive modulation such as focusing on versus distracting oneself from pain (Woo et al., 2015). Direct stimulation of any single one of these regions seldom evokes pain in humans (Penfield and Boldrey, 1937, Penfield, 1968). Mazzola and colleagues analyzed behavioral responses to electrical cortical stimulation by more than 4,000 intracerebral electrodes in 164 individuals (Mazzola et al., 2012). The only region where stimulation caused pain was in the medial parietal operculum and neighboring posterior insula. Furthermore, electrically induced pain responses in this region show a certain somatotopic spatial organization along the rostro-caudal axis (Mazzola et al., 2009).

For many years, the mesolimbic dopaminergic system, involving the ventral tegmentum, amygdala, and ventral striatum, was assumed to be responsible for pleasure processing. This idea grew from observations that rodents with microelectrode implants in mesolimbic locations (e.g. nucleus accumbens) would self-stimulate to obtain electrical stimulation from the electrode (Olds and Milner, 1954, Valenstein et al., 1970, Shizgal et al., 2001). By turning the current on only at certain locations in the testing apparatus, the animals would return repeatedly to this location, sometimes preferring this location to locations where food was provided. When given the ability to turn on the current themselves by pulling a lever, they would obsessively pull the lever – sometimes up to 2,000 times per hour (Olds, 1956).

Some similar experiments were even performed in human patients with mental illnesses. These patients engaged sometimes obsessively in “lever pressing”, which released electrical pulses from electrode implants in various subcortical locations (Heath, 1972, Portenoy et al., 1986). However, while their behavior certainly involved excessive “wanting”, it is not clear whether they actually enjoyed these pulses (Berridge and Kringelbach, 2008, Green et al., 2010, Smith, 2010, Kringelbach and Berridge, 2012).

Blockade of dopaminergic signaling typically disrupts reward-directed and consummatory behavior in rodents (Berridge and Robinson, 1998, Schultz, 2002). Extensive destruction of dopaminergic neurons can completely abolish rats’ interest in food, to the extent that they will starve to death unless artificially fed (Berridge and Robinson, 1998). Through molecular imaging in humans, dopamine signaling has been associated with a wide range of reward-related activities (Egerton et al., 2009), e.g. anticipation and emotional reactions to pleasurable music, presentation of cocaine, drug-associated stimuli, video games,

and monetary rewards (Breiter et al., 1997, Volkow et al., 1997, Koeppe et al., 1998, Scott et al., 2007, Salimpoor et al., 2011).

These findings led to the widespread idea of dopamine as a “common neural currency” for pleasant rewards (Schoultz, 2002). However, while manipulations of dopaminergic signaling or microinjections of dopamine into different parts of this “reward network” often increase how much the animal would work to obtain a reward, it does little to change the hedonic impact of the reward, as measured by how much animals lick their lips when consuming sucrose (Figure 3.1b). Although stimulation of these “pleasure centers” increases wanting of food rewards, it does not enhance “liking”, according to simultaneous measures of objective facial responses (Berridge and Valenstein, 1991).

In contrast, microinjections of opioids and certain other neurochemicals into discrete locations in ventral pallidum and the rostral part of the NAc enhance the intensity of actual “liking” (Pecina and Berridge, 2005). The “hedonic hotspots”, where microinjections of opioids increase “liking” are very small in size compared to locations where opioid or dopamine microinjections increase “wanting” responses (Figure 3.1a). The hotspots correspond to <10% of the accumbens shell (about one mm³ in rodents and one cm³ in humans, if proportional) and the ventral pallidum. The hotspots in the NAc shell and ventral pallidum seem to work as a single integrated circuit in the rat brain. Opioid stimulation in the NAc hotspot increases liking reactions to sucrose and corresponding firing signals in the ventral pallidum hotspot (Smith et al., 2011). Further, pharmacological opioid blockage of either of these hotspots prevents amplification of “liking” by activation of the other, suggesting that cooperation of both hotspots is needed for improving hedonic impact (Smith and Berridge, 2007). Importantly, however, stimulation of these hotspots with antagonist drugs such as naloxone has not been found to *abolish* “liking” responses to palatable tastes. Additional potential hotspots for “liking” reactions have been discovered in rodents in a prefrontal region that may correspond to Orbitofrontal Cortex (OFC) in humans, and in the rat homologue for the human insula (Castro et al., 2014).

Although the hedonic “hotspots” appear to be few and cover only a small fraction of brain regions that generate reward-related behaviors, liking circuitry appears resilient. Few, if any human cases are known, where basic liking reactions are completely abolished. One case report describes a patient whose life became overshadowed by anhedonia and depression, with much reduced positive affect and attenuated craving for rewards. These symptoms appeared after selective damage to bilateral ventral pallidum, which was performed to alleviate Parkinsonian symptoms (Miller et al., 2006).

Within the accumbens shell, μ -opioids increase liking only when microinjected within the small rostroventral hotspot. When microinjected in nearby locations, the same substance instead increases “wanting” without “liking”, suppresses aversive “disgust” reactions, or suppresses both “liking” and “disliking” reactions to sweet or bitter tastes (Figure 3.2). Moreover, there is recent evidence that other substances, like orexin and cannabinoids, may have comparable effects within these hotspots (Mahler et al., 2007, Ho and Berridge, 2013).

This may reflect a strong flexibility of this circuit, allowing other pharmacological agents to employ the role of opioids if this system is disrupted.

The central role of the μ -opioid system for reward in rodents appears to be preserved across mammalian species. μ -opioids mediate preference for the most valuable option when several rewards are available. For instance, blockade of μ -opioid signaling using antagonist treatment reduced consumption of palatable cookies in rats but did not affect how much standard chow was eaten (Cooper and Turkish, 1989). Conversely, enhancement of μ -opioid signaling with agonist treatment specifically enhanced male rats' display of sexual "wanting" of female rats in estrus, but not not-estrus (Mahler and Berridge, 2012). A parallel finding in humans showed that psychopharmacological treatment with μ -opioid agonism increased, while antagonism reduced, the relative appeal of highly attractive over less attractive opposite-sex faces, reflected in both rated attractiveness ("liking") and desire to keep viewing ("wanting") (Chelnokova et al., 2014).

In rodents, AMPA-blocking or GABA-stimulating microinjection within the NAc shell can produce a range of positive and negative affective responses depending on location, resembling an "affective keyboard" (Faure et al., 2008). While microinjections in rostral locations generate eating responses in the animals, injections of the same drug in more caudal locations instead induce displays of disgust or fearful behavior (Richard et al., 2013). A similar negative-to-positive affective "gradient" has been suggested for the orbito-frontal cortex, along the medio-lateral axis, based on human fMRI studies reporting representations of positive hedonic processing more medially and

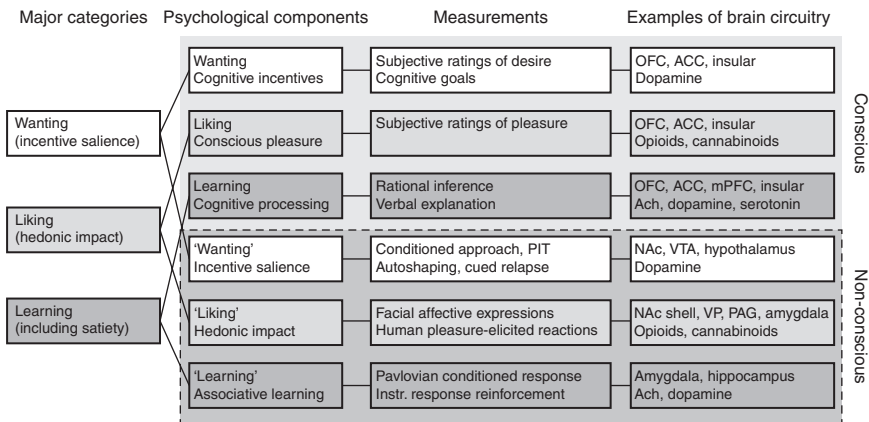


Figure 3.2 Measuring hedonic value. Hedonic processes may be divided into at least three major neurobiological and psychological components: wanting or incentive salience (white), liking or hedonic impact (light gray), and learning (gray). These components have conscious aspects, which can only be investigated in humans, and non-conscious aspects, which can also be investigated in non-human animals. Types of measurements are listed in the second column, and examples of brain circuitry involved in each subcomponent are listed in the third column.

Source: Borrowed Figure 2 (page 5) from Berridge, K. C. and M. L. Kringelbach (2011).

negative hedonics more laterally (Kringelbach, 2005). Two neuroimaging studies in humans have reported affective gradients within the ventral striatum (Seymour et al., 2007, Baliki et al., 2013). While not all of these studies address pain specifically, the findings may point more generally to how core pleasure and displeasure are generated, which may have implications for the affective aspect of pain.

3 The malleability of pain and pleasure: the importance of utility

In a rapidly changing world, being able to differentiate between typical “reward” signals like high-calorie food and a “punishment” signal like tissue damage is not enough to ensure survival and procreation. In order to make adaptive decisions on how to act in a given context (in response to a given set of stimuli), it is necessary to be able to prioritize these signals in light of surrounding sensory stimuli, current internal (interoceptive) state, and previous knowledge. Bayesian decision analyses have provided evidence that sensory systems use internally available information (“priors”) to make proactive inferences about sensory signals, and that this is updated by learning (Knill and Richard, 1996, Friston and Kiebel, 2009). Such “priors” include external (other sensory information) and internal (such as affective-motivational state) information, in light of existing (learned) knowledge of what such information means. This not only affects behavioral responses, but also shapes the actual sensory experience (Kersten et al., 2004, Fletcher and Frith, 2009, Schmack et al., 2013). In other words, sensory systems work more like active interpreters of meaning than as passive measuring instruments of the environment. Importantly, this is not only the case for sensory pleasure/displeasure, but may reflect a general mechanism of sensory systems, encompassing also more non-affective discriminative aspects of sensations. A general assumption made by such Bayesian accounts is that sensory systems in general are not *primarily* aiming to convey sensory representations that match, as closely as possible, the source of the stimulus (the environment or bodily state), but rather to decide which sensation should be had, given the information at hand (Friston, 2005, Kveraga et al., 2007). This does not mean sensory systems are not engaged in producing accurate and reliable estimations of external objects. In many circumstances, it is arguably favorable to have accurate and consistent sensations, such as when visually perceiving the letters in this text, which likely appear to have shapes and colors regardless of prior information. However, when stimuli are more ambiguous or conflicting, sensations may be affected by “priors” to a larger degree. One basic but illustrative example is the McGurk illusion, in which the auditory sensation when listening to a person saying “ba ba” is altered by viewing a person saying something slightly different, e.g. “ga ga” (McGurk and MacDonald, 1976). Rather than having a conflicting sensation of seeing a person mimicking “ga ga” while hearing a person saying “ba ba”, a majority of people will have a coherent audiovisual sensation of a person saying “ga ga” or “da da”. In this case, which likely doesn’t involve much hedonic

processing, the visual information of lip movements is used to interpret the temporally synced auditory stimuli (and perhaps also the other way around).

The flexibility and constraints of top-down influences on sensation are still being explored, but it is likely that for the perception of most stimuli, the (un)pleasantness is more malleable than non-affective characteristics. In fact, very few (if any) reward stimuli are ubiquitously pleasurable. Even stimuli that seem inherently “positive”, like sweet taste, can become hedonically “flipped”. Eating a delicious chocolate can be intensely pleasant if you are hungry for chocolate. Yet the delight turns to disgust if you keep eating it beyond satiety, although the sensory stimulus remains the same (Small et al., 2001). Similarly, while a hot bath is likely very pleasant if you just came in freezing from a winter storm, you may prefer an invigorating cold shower if you’re boiling in the midst of a heat wave.

Introducing the concept of alliesthesia, Cabanac (1971) postulated that stimuli that serve to move the organism towards physiological or psychological homeostasis should be perceived as pleasant, while stimuli that serve to move the organism out of homeostasis should be perceived as unpleasant or painful. Rather than the inherent quality of the stimulus itself, it is the *utility* of the stimulus for the organism at the time that determines the hedonic value of the stimulus. The taste of salt at concentrations higher than seawater is usually unpleasant to humans and causes “disliking” gape reactions in rats. However, if sodium levels are experimentally depleted, which induces a state of “salt appetite”, rats will instead display “liking” reactions comparable to that of sucrose taste (Berridge et al., 1984, Tindell et al., 2006). Moreover, this “hedonic flip” is mirrored by the firing rate of neurons in the ventral pallidum, indicating that coding of basic sensory pleasure depends here on the homeostatic utility, and not merely intrinsic qualities, of the stimulus (Tindell et al., 2006).

Similarly, the often positively laden experiences of eating spicy food, burning muscle during physical exercise, or engaging in endurance sports illustrate that pain does not always carry an exclusively negative hedonic value. While muscle pain is inevitable during a marathon run, it may nevertheless carry a certain positive value, if the pain represents a challenge rather than a threat. This change in meaning can reduce the negative hedonic feeling usually associated with pain, as well as the behavioral response to it. A recent study assessed the impact of positive or negative verbal instructions on the tolerance of ischemic arm pain. Participants who were told that the pain would have beneficial effects on muscle cells tolerated the pain for almost twice as long as participants who were informed only about the aversive aspects of the pain (Benedetti et al., 2013) (Figure 3.3b). Moreover, moderate pain can be “hedonically flipped” from unpleasant to pleasant if the pain represents safety from an even more painful stimulus (Leknes et al., 2013) (Figure 3.3a). This hedonic flip is underpinned by increased co-activation of the PAG with NAc and ventromedial Prefrontal Cortex (vmPFC) (Leknes et al., 2013). Pain can also enhance the pleasure of reward hedonics. For example, people report greater pleasure from chocolate after undergoing a cold-pressor test, where the task is to keep one’s hand dipped in a bath of cold water ($\sim 1^{\circ}\text{C}$) for as long as possible (Bastian et al., 2014a).

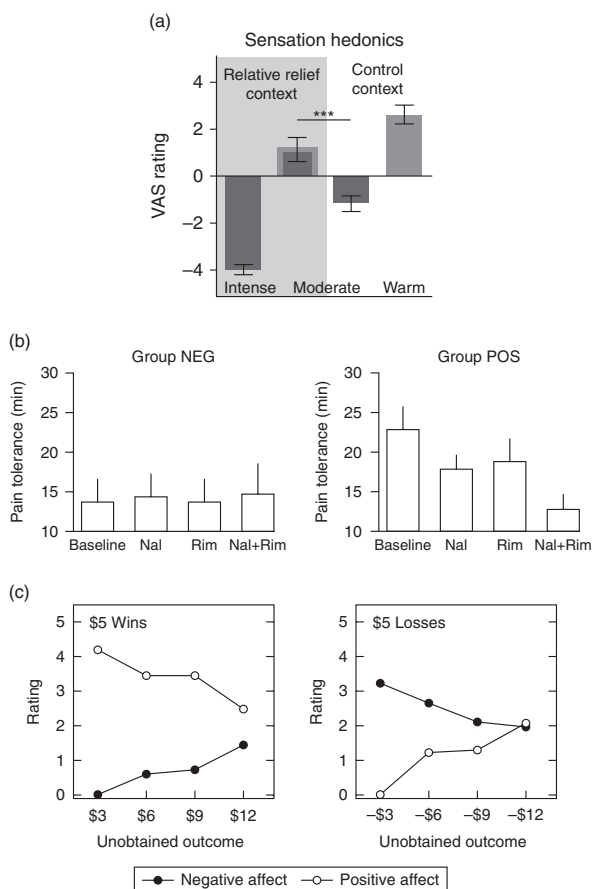


Figure 3.3 Contextual modulation of hedonic value. (a) In a recent study, the occurrence of a moderately painful stimulus represented either the best possible outcome (avoidance of intense pain – relief context) or the worst possible outcome (moderate pain instead of mild non-painful heat – control context). Whereas the moderately painful heat stimulus was rated as unpleasant in the control context, the same pain stimulus was rated as pleasant when it represented safety from an even worse pain (relief context). (b) In another experimental study, one group of participants were told that experimental muscle pain would hurt but had therapeutic benefits (POS), while another group were only told that pain would hurt (NEG). Pain tolerance in the (POS) group was almost doubled compared to the (NEG) group. This increase in pain tolerance was partially reversed when participants were pre-treated with either naloxone (Nal) or the endocannabinoid receptor antagonist rimonabant (Rim), and completely abolished when pre-treated with both Nal and Rim. (c) In a study investigating affective responses to monetary gains or losses under different contexts, it was found that winning \$5 was associated with negative affect and reduced positive affect when it was possible to win even more, while losing \$5 was associated with positive affect and reduced negative affect when the alternative was to lose even more. In fact, *winning* \$5 instead of \$12 induced almost the same level of positive and negative affect as *losing* \$5 instead of \$12.

Source: (a) Borrowed figure 3 (page 406) from Leknes, S. et al. (2013). (b) Borrowed figure 3 (page 365) from Benedetti, F. et al. (2013). (c) Borrowed figure 2 (page 327) from Larsen, J. T. et al. (2004).

A possible mechanism whereby painful experiences can be pleasant, e.g. in consumption of spicy food, is that the enhanced pain-induced arousal increases attention also to positive sensory signals, resulting in a more intense, and also pleasant, sensory experience (Bastian et al., 2014b, Leknes and Bastian, 2014). Another possibility is that endogenous release of endorphins caused by aversive signaling accounts for the positive affect (Carstens et al., 2002). Molecular imaging and placebo studies employing pharmacological μ -opioid blockade have shown that endogenous μ -opioid signaling contributes to downregulation of pain (Levine et al., 1978, Zubieta et al., 2001, Zubieta et al., 2005, Eippert et al., 2009a). However, one study found that offset analgesia, the disproportionately large analgesic response that often follows a slight reduction in the intensity of an ongoing pain, was unaffected by administration of an opioid antagonist or an agonist (Martucci et al., 2012).

4 Hedonic value is modulated by motivational states

The hedonic valuation of rewards and punishments are both influenced by the overarching motivational state of the organism (Loeth et al., 2014). The conceptual meaning, and the corresponding pleasure/displeasure attributed to a stimulus, depends on the predicted role of this stimulus in achieving the goal currently kept in mind. The prediction of outcome value, or cost, of different actions is a constantly ongoing process and is reflected in hedonic experience (Friston and Kiebel, 2009, O'Reilly et al., 2013). The motivation-decision model of pain, as proposed by Fields (Fields, 2006, 2007), describes brain mechanisms that enhance or reduce the hedonic impact of events based on their relative importance at a given time. The model was initially put forward to explain modulation of pain, but the basic idea holds for all events that fall within a reward-punishment continuum. Fields postulates that – as a result of an unconscious decision-making process – any concurrent or impending event that is more important for the individual than a pain stimulus should suppress the hedonic impact of this pain. The event of superior importance may for instance be a greater threat or a potential reward. Likewise, anything judged as more important than an impending reward – for example, a threat or a bigger reward – should suppress the hedonic impact of this reward.

A central question in affective neuroscience is how the brain modulates this hedonic impact. Does it target the neural systems that generate pleasure or displeasure, e.g. hedonic hot- and coldspots, or does it also modulate ascending sensory information that gives rise to pleasure or displeasure? If so, at which levels does this modulation take place? There is well-established evidence that ascending nociceptive neurons in the spinal dorsal horn are modulated by the brain (Wall, 1967, Woolf, 2011). The PAG in the midbrain controls incoming nociceptive signals indirectly through the rostroventral medulla (RVM) (Millan, 2002, Fields, 2004). Neurons in the RVM project to the spinal dorsal horn, with inhibitory (“OFF cells”) or excitatory (“ON cells”) effects on nociceptive transmission (Urban and Gebhart, 1999, Neubert et al., 2004). The PAG receives

direct input from the limbic structures' amygdala and ventral striatum, and from the prefrontal cortex, constituting a pathway by which affective or cognitive information can influence ascending sensory information already at the spinal dorsal horn (Fields, 2004). Thus, by the time nociceptive signals reach the brain, they have already been processed, likely affecting the unpleasantness along with other stimulus characteristics of pain.

While much of the circuitry involved in valuation and reward is remarkably similar between humans and other mammals (Murray et al., 2007, Haber and Knutson, 2010), there are some notable differences. For example, humans have massively expanded prefrontal cortices, reflecting greater encephalization. This is likely to be accompanied by more prefrontal influence of other brain processing via increased prefrontal connectivity to other cortical and subcortical brain regions. Compared with rats, primate orbitofrontal cortex projects more clearly defined descending connections to the hypothalamus and brainstem structures. Thus, "higher-order" cognitive and affective information may have more influence over hedonic value in humans compared to other animals. This should be the case for both positive and negative hedonic experiences, and indeed evidence of prefrontal regulation of pain is rife in the literature (e.g. Lorenz et al., 2003, Seminowicz et al., 2011, Taylor et al., 2012).

Functional neuroimaging in humans suggests that orbitofrontal, insular, and ventromedial prefrontal cortices (as well as the amygdala, PAG, and VTA) play important roles in reward processing. Activity in this circuitry often correlates with self-reported pleasantness of e.g. food, touch, sexual orgasms, drugs, chocolate, and music (Blood and Zatorre, 2001, Small et al., 2001, Rolls et al., 2003, Vollm et al., 2004, Kringelbach, 2005, Georgiadis et al., 2006, Grabenhorst et al., 2008, McGlone et al., 2012, Ellingsen et al., 2013). Nevertheless, while large-scale disruption of these regions often has consequences for the aforementioned processes, it seldom abolishes the capacity for normal pleasure and displeasure. Thousands of human patients received prefrontal lobotomies in the 1950s with massive damage to the ACC and OFC. However, in spite of clear deficits in decision-making and dramatic personality changes, these patients did not seem to lose the capacity for hedonic feelings and continued to live affective lives (Valenstein, 1986, Damasio, 2000). Further, case reports suggest that affective processing may be relatively intact in patients with insular damage. One patient had bilateral insular cortices completely destroyed from Herpes simplex encephalitis. Although he showed massive learning deficits and was unable to remember any new factual item for more than 45 seconds, he still showed differential preference for food and reported feelings of happiness, pleasure, and pain (Damasio et al., 2013). Another case report of patients with insular damage reported intact feelings of pain, with increased pain intensity for experimental pain stimuli (Starr et al., 2009), in line with a role for the insula in pain regulation (Craig, 2009). Lastly, a case report of a 6-year-old patient with hydrocephalus, with the absence of cerebral tissue rostral to the thalamus, reported that the boy "smiled when spoken to and giggled when played with" (Shewmon et al., 1999, p. 366). Together, these cases suggest that although

playing important roles in the hedonic valuation, cortical nodes of hedonic processing are not necessary for hedonic experience. Instead, these regions are likely to represent hedonic value associated with other processes, such as learning and memory, cognitive representations, language, decisions, action, or consciousness (Kringelbach and Berridge, 2010, Berridge and Kringelbach, 2013).

5 Modulation of hedonic experience by expectations, learning, and contextual meaning

A range of neuroimaging studies in humans show that cognitive or affective modulation of pain can alter widespread somatosensory processing in the brain (Wager et al., 2004, Tracey and Mantyh, 2007, Berna et al., 2010, Knudsen et al., 2011, Amanzio et al., 2013, Ellingsen et al., 2013). A useful experimental model for probing top-down modulation of pain, but also hedonic value in general, is placebo responses, i.e. positive treatment outcomes that are not caused by the physical properties of this treatment but by the meaning ascribed to it. Although traditionally considered mostly a nuisance, or bias, that needed to be controlled for in clinical trials, the scientific interest in placebo effects has boomed in recent years, together with the development of neuroimaging tools that have provided new ways of probing the underlying neurobiology. Placebo effects are generally assumed to arise from explicit or implicit expectations, or predictions, of positive treatment outcomes (Finniss et al., 2010, Benedetti, 2014). This prediction may involve explicit expectation of a specific outcome, or arise from simple associative learning, like classical conditioning, without much explicit cognition (Jensen et al., 2012, Jensen et al., 2015). In both cases, learning plays a key role. Recent accounts focus on placebo effects as a mechanism of reward prediction, where the sensory cues associated with clinical care (e.g. the doctor's white coat, reassuring words, the sensory feedback of the treatment itself) become learned safety cues (Fields, 2004, Benedetti, 2013, Schmidt et al., 2014). In other words, the patient has learned that these cues predict relief. Along these lines, Bayesian decision theory has been employed to explain placebo analgesia as a perceptual decision, suggesting that expectation based on available sensory information and internal knowledge differentially shapes the pain experience depending on the certainty of the expected outcome (Buchel et al., 2014). Specifically, given a specific predicted outcome (e.g. a pain experience of 40 on a 0–100 scale), a high certainty of this outcome would exert a strong influence on pain (i.e. placebo analgesia if the expected stimulus is milder than the actual outcome; nocebo hyperalgesia if the expected stimulus is more intense than the actual outcome) (Anchisi and Zanon, 2015). However, such certain predictions are also associated with a high sensitivity of detecting a violation of the prediction, if the actual outcome (i.e. nociceptive signal) does not match what was predicted. This would lead to extinction of placebo analgesia. If the certainty is low, however, the prediction has a less strong influence on pain (i.e. less robust placebo analgesia) but is also less sensitive to a violation of the prediction. A recent experimental study tested this hypothesis and found that people who learned, through conditioning, that “activation” of

a sham Transcutaneous Electrical Nerve Stimulation (TENS) led to pain relief with 100% certainty, showed stronger initial placebo analgesia but with faster extinction, compared with people who learned that the cue led to pain relief with a 62.5% certainty (Au Yeung et al., 2014).

Recent advances using neuroimaging have provided insight into the neural networks responsible for placebo improvement. In line with the view of the brain as an active interpreter/predictor, there is growing evidence that placebo improvement (at least in some cases) is underpinned by top-down modulation of neural circuitry that traditionally are considered pathways for bottom-up, ascending sensory signals. In these studies, placebo-induced reduction of pain is often underpinned by widespread reductions of somatosensory (pain) processing in thalamus, insula, primary and secondary somatosensory areas, and dorsal ACC (Eippert et al., 2009a, Lu et al., 2010, Amanzio et al., 2013). Further, activity in a “pain modulatory network” consisting of ventromedial PFC, OFC, ventral striatum, amygdala, and the midbrain, is instead increased, and is thought to be responsible for the suppression of pain processing. This modulatory network is dependent on opioid processing, since administration of the opioid receptor antagonist naloxone can attenuate both the reduced pain reports (Levine et al., 1978, Amanzio and Benedetti, 1999) and the reduction in pain-related processing (Eippert et al., 2009a). The functional architecture of modulatory networks for placebo responsiveness has yet to be disentangled, but as mentioned, these regions play central roles in hedonic valuation more generally, in the monitoring and updating of expectation and the integration of available relevant information. Outside the domain of pain, Petrovic and colleagues (2005) used a conditioning paradigm, whereby participants were shown threatening images before and after administration of the anxiolytic drug midazolam. This drug robustly reduced self-reported unpleasantness from viewing the images. In a subsequent session, participants who were given a placebo labeled as midazolam reported reductions in unpleasantness comparable to the active substance. Furthermore, the placebo improvement was underpinned by increased fMRI activation in ventral striatum, rostral ACC, and mid-lateral OFC, but suppressed activation in visual cortex.

Recent findings suggest that nociceptive processing at the spinal dorsal horn can be altered when people expect pain relief. Eippert et al. (Eippert et al., 2009b) recorded fMRI of the spinal segment receiving information from the arm, and found increased activity when people had received placebo treatment that also reduced their arm pain. Similarly, attentional modulation of nociceptive signals in the spinal cord have been found when people are performing a cognitively demanding working memory task, which induces analgesia (Sprenger et al., 2012). Within the framework of the motivation-decision model, this similarity between expectancy- and distraction-induced analgesia is interesting (Atlas et al., 2013) and may reflect a more general mechanism whereby motivational states control the presence of pain. In both cases, the importance of pain is reduced, either because the brain is convinced an “analgesic” treatment will reduce pain, or because attention is focused on another, more important task. A consequence of the brain’s decision to “respond less” to pain is that the individual can focus attention on other important tasks, for example, reward seeking.

Nocebo hyperalgesia, the worsening of pain by negative expectations, may work by a comparable mechanism (Scott et al., 2008, Rief et al., 2011). fMRI studies indicate that expectancy-induced increases in pain-related negative hedonics are underpinned by amplified somatosensory responses to pain stimuli (Kong et al., 2008, Rodriguez-Raecke et al., 2010, Bingel et al., 2011). Further, a recent study investigated spinal fMRI activity during nocebo hyperalgesia and found increased activation in the relevant spinal cord segment when people expected pain to be increased (Geuter and Buchel, 2013). These studies indicate that psychological processes, in this case expectation of treatment benefit, are able to modulate sensory information along the entire sensory neural “axis”, stretching from sensory circuitry in the brain to coupling stations in the spinal dorsal horn, resulting in this context in reduced or amplified pain experience. Such modulation is less studied for non-nociceptive sensory processing or positive hedonic experiences, although work from the Berridge lab established a role of brainstem neurons in the parabrachial nucleus in the amplification of “liking” face reactivity measures to tastes (Soderpalm and Berridge, 2000). It may be that hedonic reactions to tastes rely on a signaling pathway comparable to the better-characterized pain neuraxis and that hedonic hot- and coldspots regulate incoming taste signals via the parabrachial nucleus.

If boosting the pleasure of a pleasant sensation (hyperhedonia) works in a corresponding manner, we would expect sensory activity of this appetitive stimulus instead to be increased. We recently found that a placebo modulation boosting the pleasantness of gentle touch might rely on a mechanism similar to that of placebo analgesia (Ellingsen et al., 2013). We suggested to a group of healthy volunteers that a nasal spray would increase both the pleasantness of gentle touch and reduce the unpleasantness of pain. After self-administration of this nasal spray, a placebo, people reported increased touch pleasantness and reduced pain unpleasantness. The reported improvements of pleasure and pain were proportional to their prior expectations of treatment effects. While fMRI recordings during pain showed decreased somatosensory processing, recordings during gentle touch stimuli showed instead amplified activity in somatosensory areas (SI, SII, insula). Moreover, the magnitude of both hyperhedonic increases and analgesic decreases in somatosensory responses were accounted for by the individual strength of the functional coupling between medial OFC and PAG. It remains to be seen whether the modulation of pleasurable touch, like pain, involves descending modulation of cutaneous afferents in the spinal cord, perhaps via the RVM. It will also be interesting to see whether this modulation relies on opioid processing. These results cannot tell us the direction of causality. Nevertheless, this finding may suggest that, like negative hedonics such as pain, psychological modulation of pleasant sensations involves a more comprehensive modulation of the underlying sensory processing, and not only within higher-level valuation circuitry.

Many of the ingredients of the placebo effect, like subtle associative learning of safety signals, and the therapeutic ritual of being cared for by a (professional) conspecific, may build on relatively generalized mechanisms that stretch far outside the context of medical treatment, e.g. explaining phenomena such as people

who feel and behave intoxicated after drinking non-alcoholic beverages (Hull and Bond, 1986, Knight et al., 1986) or otherwise behaving according to expectations even when these are in violation of sensory evidence. They may also overlap with not only a range of alternative and complementary treatments but also with prominent psychological therapies (Rosenthal and Frank, 1956, Kirsch, 2005), like cognitive behavioral therapy and emotional and cognitive reappraisal strategies (Tracey, 2010, Buhle et al., 2013).

6 Mixed emotions

While sensations that carry a hedonic “gloss” are often thought of as either positive (reward) or negative (punishment), there are many examples of experiences that carry both positive and negative hedonic value. Such “duality” may be more frequent in composite emotional experiences, such as feelings of nostalgia, guilty pleasures, or bittersweet memories, but are also present in more basic sensations like deep pressure massage or being tickled. A pleasure-displeasure dichotomy may therefore be too simple.

In one study, people engaged in a gambling game with two types of contexts. In one context there was a 50/50 chance of winning either of two amounts of money, while in the other context there was a 50/50 chance of losing either of two amounts (Larsen et al., 2004) (Figure 3.3c). Thus, winning the lower amount was the worst possible outcome in the first context, while losing the lower amount was the best possible outcome in the second context. Participants reported feeling both good and bad about the disappointing wins and relieving losses. Moreover, when participants indicated with two buttons whenever they felt good or bad, participants reported feeling good or bad simultaneously rather than alternating between the two, suggesting that these constituted unitary mixed emotional events. Nevertheless, since the positive and negative feelings in this study were attached to separable aspects of this emotional experience – for disappointing wins, winning money (positive) but winning the lower amount (negative) – it is possible that these are more akin to separable experiences “lumped” into a composite emotional event, in a similar way that the redness and blueness in the British flag could be both separable sensations and parts of a unitary composite sensation of the Union Jack. It is not currently known whether, if one were to deconstruct sensations to the smallest constituent unit, it is possible for such a unit to have both positive and negative hedonic value.

The relative value of wins and losses is encoded in primate OFC neurons, where the same neurons have been shown to track the best option in a given context, regardless of whether the option is the larger of two wins or the smaller of two losses (Hosokawa et al., 2007). These signals may correspond not to the experience of mixed emotions, but instead to a related process, namely the overall value of the event. In other words, although animals may experience concurrent pleasure and pain, it is nevertheless possible to compute an overall value that helps answer an important question: Should one repeat the actions that led to the experience?

Rozin and colleagues refer to the enjoyment of experiences that are innately aversive, but do not represent a threat, as “benign masochism” (Rozin and Schiller, 1980, Rozin et al., 2013). According to this idea, when the brain judges such an aversive sensation to represent no real danger, then a sense of control, or mastery, gives rise to positive affect.

There is little evidence for the preference of mixed emotions in animals. Rats display both aversive and “liking” taste reactivity responses to bittersweet foods (Doyle et al., 1993), but would likely prefer a purely sweet substance over the bittersweet. Typically a stimulus is associated with either avoidance or approach behavior by the animal, not both at the same time (Rozin et al., 1979; however, see Rozin and Kennel, 1983). This is in line with the notion that hedonic impact may be more influenced by top-down factors in humans compared other mammals (Berridge and Kringelbach, 2013).

Although most of the studies investigating mixed emotions address explicit hedonic feelings, either through here-and-now assessments or in retrospect, they tell relatively little about how this liking or disliking relate to motivational or learning aspects of rewards. For example, although an experience can be simultaneously liked and disliked, it could still be that motivational processes follow a more either/or principle – that an event cannot be both worked for and against at the same time. More research is needed to definitely answer whether it is meaningful to classify sensations on a bipolar continuum with pleasure and displeasure as endpoints, or whether they are separable processes that compete for preference in the brain but are not mutually exclusive. Nevertheless, it seems clear that at least humans are capable of experiencing unitary feelings that are pleasant and unpleasant at the same time.

7 Conclusions

We have reviewed the neural basis of hedonic processing involved in pleasure and pain in humans and animals. Extensive evidence suggests that hedonic events can be functionally and neuroanatomically segregated into (dis)liking, motivational, and learning components. The hedonic value of both basic sensory and “higher-order” pleasure and displeasure is often strongly influenced by multisensory contextual information, as well as the current short- and long-term needs and goals of the organism in terms primarily of survival. Recent investigations of placebo effects have revealed that such conceptual information is not only encoded in higher-order brain circuitry for valuation but can actively modify “sensory” circuitry traditionally thought of as “labeled” pathways for ascending sensory signals. Rather than being passive “windows to the world”, sensory systems perpetually make active, predictive inferences about the usefulness and value of sensory information. Consequently, sensations may reflect unconscious decisions about which sensation *should be had*, given the information at hand, which in turn may help select the most appropriate action.

Part II

Pain and emotion

4 The rationality of emotional and physical suffering

Michael Brady

Negative emotions can sometimes be reasonable, sometimes unreasonable. My anger at the Principal's pay rise is reasonable, given that my colleagues and I are facing pay cuts. My fear when the plane hits mild turbulence is unreasonable, since I know that mild turbulence is not dangerous. Physical pain, on the other hand, cannot be reasonable or unreasonable. The pain in my toe, when some clodhopper on the dance floor steps on my foot, is a *causal* response to clumsiness; but it is not a rational, reasonable, warranted, or justified reaction. The chronic pain in my back, for which the doctors can find no evidence of injury, no doubt has *some* cause; but it isn't an irrational, unreasonable, unwarranted, or unjustified reaction to that cause. So, at least, common sense tells us.

In this paper I want to support this common-sense intuition and attempt to explain why there is this purported *normative* difference between negative emotions and physical pain. In Section 1 I'll argue that there are good reasons to think that emotional suffering is reason-responsive, because emotional suffering is an essentially evaluative phenomenon. In Section 2 I'll show that none of these are reasons to think that physical suffering is an essentially evaluative phenomenon, and that we have, as a result, no good reason to think that it is reason-responsive. And in Section 3 I'll close by showing how this puts pressure on a new and upcoming theory in the philosophy of pain, namely *evaluativism*.

1

In order to keep things simple, I want to focus on the putative normative difference between cases of emotional suffering – such as grief, shame, disappointment, guilt, despair, anxiety, loneliness, lovesickness, and the like – and cases of physical suffering – paradigmatically pain, but also including experiences of coldness, tiredness, hunger, nausea, and irritation. In this section I'll try to make the case that emotional suffering can be reasonable or unreasonable; in the section to follow I'll cast doubt upon whether anything similar can be said about pain and other kinds of physical discomfort.

The common-sense intuition that emotional suffering can be reasonable or unreasonable is best supported, in my view, by looking more closely at a platitude about the *nature* of emotion. In "Feelings that Matter", Annette Baier writes:

“[w]e all accept the idea that emotions are reactions to matters of apparent importance to us: fear to danger, surprise to the unexpected, outrage to insult, disgust to what will make us sick, envy of the more favoured, gratitude for benefactors, hate for enemies, love for friends, and so on.”¹ Emotions, pretty much everyone agrees, are reactions to a class of real or apparent *values*, broadly construed. So fear is a response to danger, anger to insult, guilt to moral fault, shame to what is shameful, pride to achievement, and grief to loss. If so, however, then we have good grounds for thinking that emotional reactions are *normative* rather than merely *causal* responses to values or disvalues. This is because the values themselves are best understood as things that normatively require, rather than simply tend to cause, the reactions in question.²

To see this, we can note something else about which there is widespread agreement: that concepts and/or properties like danger, shameful, moral fault, loss, and the like have *some* connection with emotional reactions of fear, shame, guilt, grief, and so on. It would be bizarre to claim that shameful, for instance, has *nothing* to do with feelings of shame. Significant disputes arise, of course, as to the precise nature of this relationship. One possibility is that some X is shameful iff a subject S is ashamed of X. Here we understand the concept or property in terms of emotional responses *actually* caused or elicited. This causal view is highly implausible, however, since there can clearly be divergence between something’s being shameful, for instance, and our feeling ashamed. Perhaps I’m ashamed of things, such as my northern accent, that aren’t shameful. Perhaps I’ve behaved shamefully at the party but am not ashamed, because I can’t remember what I did. So we should reject any such strong link between concept/property and response.

A more plausible view is *dispositionalism* about value. On this account, divergence between (to pick a different property-emotion pair) moral wrongdoing and guilt is possible. Nevertheless, they are intimately connected, since moral wrongdoing *tends to cause* guilt in normal subjects under normal conditions.³ Values, on this view, are akin to colours: redness, on an analogous dispositionalist account, is simply that property which tends to cause a visual experience of redness in normal subjects under normal conditions. Divergence between evaluative property and response – as when one is aware that one is in danger but isn’t afraid, or when one feels guilty but hasn’t done anything wrong – is explained by appeal to some way in which conditions are abnormal, or some non-rational fault in the subject. To take the parallel case in colour perception: if I don’t see the postbox as red, then either conditions are abnormal (perhaps there is a blue light shining on it) or there is some non-rational fault in me (perhaps something has gone wrong with my eyes).⁴

Dispositional accounts of value are an improvement on basic causal accounts. However, such accounts face a devastating objection, at least when applied to value: for such accounts fail to allow that value can clearly diverge from what can be regarded as normal responses in normal conditions.⁵ To see this, note that there can be significant divergence between value and emotional response if we appeal to *statistical* or *social* norms. I imagine that in the 1930s it was normal – in the sense of widely accepted, and in that sense as a part of the moral and social

fabric of British culture – to think that homosexuality was shameful. So homosexuality would tend to elicit feelings or judgements of shame, in normal people in normal conditions, in 1930s Britain. But it clearly doesn't follow that homosexuality was shameful then, any more than now. The dispositionalist will therefore have to tailor the meaning of "normal" so that the theory doesn't imply that homosexuality was or is shameful. But how are they to do this? One possibility would be to appeal to some response-independent standard to rule out the majority view; but this, of course, would be simply to give up the dispositionalist project of trying to explain concepts and properties in terms of human emotional responses. Another possibility would be to argue that most people in 1930s Britain were subject to certain biases which undermine their emotions and judgements; but this, too, would seem to involve giving up dispositionalism, at least if we think that showing that people are subject to biases is a matter of showing that they are to that extent *irrational*. If we favour the latter view, we would no longer embrace dispositionalism, but a view that is much more plausible – as I'll now explain.

A better account of evaluative concepts and properties is what Justin D'Arms and Daniel Jacobson term *rational sentimentalism*. On this view, the connection between some evaluative concept/property and an emotional reaction is not causal, nor dispositional, but *normative*. D'Arms writes:

[T]he dominant contemporary sentimentalist approach to the relationship between sentiments and evaluative concepts has not been dispositionalism, but a second-order sentimentalism, holding that to apply a response-dependent concept F to an object X (i.e. to think that X is F) is to think it appropriate (merited, rational, justified, warranted) to feel an associated sentiment F toward X.⁶

This is a view about the content of normative or evaluative thoughts or judgements. But we can easily adapt the view so that it is about the concepts or properties themselves. Thus, the rational sentimentalist will hold that the concept of "dangerous" is the concept of "meriting or warranting fear"; by the same token, they will hold that for something to *be* shameful is for it to justify or make rational the feeling of shame. Evaluative judgements or thoughts are therefore judgements or thoughts employing these concepts or about these properties.

Rational sentimentalism is not without its own problems. Sentimentalists struggle with the thorny problem of specifying the "right kind of reason" for some relevant emotional response.⁷ Nevertheless, the position marks a significant improvement on dispositionalism. For one thing, it allows there to be significant divergence between some value and what is socially or statistically normal: the fact that most people in the 1930s thought homosexuality shameful has no bearing on whether homosexuality was or is shameful, since we might be confident that such attitudes failed to track any reasons to think homosexuality shameful. And we might be confident of this because we are confident that there are no such reasons; purported "justifications" for the shamefulness of homosexuality will be bogus. For another, rational sentimentalism gives us the tools to explain

(as I've just done) *why* there is divergence between some value and the statistically normal response, without appeal to any independent standard; shamefulness is still defined in terms of the emotion or sentiment of shame, only it insists that what is shameful is what merits shame, or what gives us reason to be ashamed. Finally, rational sentimentalism captures, far better than causal or dispositionalist accounts, the nature of our thinking about values. For according to rational sentimentalism, such thinking is a matter of reflecting upon, deliberating about, and identifying those considerations that constitute genuine reasons for emotional responses. But this is precisely what one tries to do, when one tries to think about whether one should feel guilty about not donating more to Oxfam, or whether one should be ashamed of one's accent, or anxious about the forthcoming lecture, or disappointed about the government policy, or angry about the research assessment results. Such thinking isn't, plausibly, thinking about what counts as a normal subject or normal conditions. It is, instead, thinking about the considerations that seem relevant to one's evaluative-emotional situation, considerations that bear on whether one rationally ought to think and feel as one does.

If rational sentimentalism is correct, then values constitute reasons for emotional responses: danger is a reason for fear, an insult a reason to be angry, a daughter's achievement a reason to feel pride, the death of a close friend a reason to grieve, a moral failing a reason to feel guilt. This follows from the *nature* of values and the related sentiments. Now, if the values in question are reasons for emotional responses, then it must be possible for a subject to have the relevant responses *for* those reasons, at least under certain conditions.⁸ It is, after all, difficult to understand the sense in which someone could have a reason for an emotional response, but under no circumstances would it be possible for them to have that emotional response for that reason. But what is it that enables emotional experience to be responsive to the reasons there are? Well, it is *generally* true that one's mental states are responsive to reasons insofar as one acknowledges or recognizes or grasps in some sense that one has a reason and forms the mental state *as a result of* this recognition or grasping or acknowledgment. Of course, the relevant recognition or grasp need not be explicit, or something of which one is consciously aware. The idea that we are rationally guided by the reasons we grasp or acknowledge does not imply that we are consciously thinking about our reasons and what they are reasons for when we act. Instead, rational guidance requires much less in the way of "foregrounded" thinking and can allow that most of the work takes place in the "background".⁹ All that might be needed, in order for one's mental states to be responsive to reasons, is the capacity to explain, after suitable reflection, the reason *why* one is in the relevant mental state. Such a capacity to explain rests, plausibly, upon the *implicit* recognition or grasp of some consideration as a reason, a recognition or grasp that can be made explicit if we are asked to justify what we believe or intend.

Now emotions can be responsive to such a recognition or grasp. Indeed, emotions are held, at least by most theorists, to be *partly constituted by* an element that serves as such a recognition or grasp. For there is widespread agreement that paradigm cases of emotion involve an appraisal or evaluation of the object or event

that is the target of that emotion. Thus fear involves an appraisal or assessment of some object or event as dangerous; anger involves an appraisal or assessment of some event as insulting or as constituting a wrong to oneself; disappointment is in part a negative evaluation of some event that violates hopeful expectation; and so on for all other forms of emotional suffering. Emotional experience is responsive to reason, in other words, because emotions are partly constituted by an evaluation or appraisal that attributes some evaluative property to an object or event. The evaluation that partly constitutes an emotional experience is thus a recognition or grasp of the fact that this very emotional experience is rationally warranted or merited. This is why, moreover, it would be irrational for one to be aware of some value and *not* have the relevant emotional response; for this is a failure to have an emotional response that is, by one's own lights, rationally merited or warranted. Once again, there is no requirement that the relevant evaluation or appraisal be explicit or consciously held. It need not consist of a propositional attitude like an evaluative judgement or belief, for instance. Still, in normal cases we are implicitly aware of the relevant value that our emotion is responsive to: we will readily cite the insulting behaviour of the driver as a reason why we are raging, or the terrible loss as a reason why we grieve. If so, then our implicit awareness of values and reasons is standardly made explicit if we are asked for or feel the need to offer a justification.

If all of this is correct, then the common-sense idea that emotions can be reasonable or unreasonable follows from a platitude about the nature of emotion, a plausible account of what values are, and an account of emotional experience as involving, as a central element, an evaluation or appraisal. On this picture, values thus provide reasons for emotional experiences, whilst the fact that emotions are evaluative enables them to respond to such reasons. In the following section I'll raise doubts as to whether anything similar can be said about pain and other forms of physical suffering.

2

Earlier I noted the common-sense view that forms of physical suffering, such as pain, hunger, coldness, and tiredness, are not reason-responsive. What I want to argue now is that such states are not reason-responsive precisely because they are not responsive to, or partly constituted by, evaluations. There is nothing in pain, hunger, coldness, or tiredness that constitutes a grasp or recognition of values or reasons, in other words; nor is there anything in the process by which we come to feel pain, cold, hunger, etc., that constitutes such a grasp or recognition. This is why these forms of physical suffering remain purely causal responses to objects and events. In what follows I'll provide three arguments for this conclusion, arguments which proceed by highlighting important differences between emotional and physical suffering. The first is an argument from introspection: we don't need to appeal to any evaluative element in order to identify pain and other kinds of physical suffering, whereas we do in the case of emotional suffering. The second is metaethical: the relevant disvalues in cases of physical suffering are not best

understood along rational sentimentalist lines, but they are, as we have seen, best understood along such lines for cases of emotional suffering. And the third is functional: emotions involve evaluations because there is a need for *flexibility* of emotional response to a very wide range of values and disvalues; but there is no such need in the case of physical responses to the causes of pain, coldness, tiredness, and hunger. Let us take these in turn.

(i) *Introspection*

For the first part of the twentieth century, the *feeling theory* of emotion held sway. Such a view, developed by William James and (independently) Carl Lange in the late nineteenth century, maintained that emotions are nothing but the consciousness of bodily changes, elicited by the perception of some object or event in our environment. As James puts things: "My thesis ... is that the bodily changes follow directly the *perception* of the exciting fact, and that our feeling of the same changes as they occur is the emotion."¹⁰ Now, a central problem with the James-Lange account, pointed out by Walter Cannon in 1929, is that the same bodily or visceral changes – such as increased heart rate, increase of blood sugar, inhibition of digestive activity, sweating, discharge of adrenalin, widening of pupils, and the like – occur in *different* emotional states, such as fear and rage, and also in *non-emotional* states, such as fever, exposure to cold, and asphyxia. Cannon concludes that "[t]he responses of the viscera seem too uniform to offer a satisfactory means of distinguishing emotions which are very different in subjective quality."¹¹ The thought here is that we cannot *differentiate* emotions purely on the basis of how they feel – we cannot tell apart fear and rage simply by focusing on the feelings of bodily change that occur in these states, or identify some emotion solely on the basis of how it feels. In order to distinguish different emotions, we need to appeal to some other elements. Cannon, Bedford, and very many philosophers since have thought that these must be *intentional* elements. In particular, emotions are to be identified and distinguished in virtue of characteristic patterns of evaluation or appraisal. So fear is distinguished from anger, not in virtue of how it feels, but because the former involves an assessment of danger whilst the latter involves an assessment of insult. The idea that we need to invoke such an evaluative structure if we are to identify particular emotions and distinguish different emotions provides a strong argument against feeling theories of the kind proposed by James and Lange; it also provides strong additional support for the view that emotions themselves are, essentially, forms of evaluative experience.¹²

A similar form of argument has little traction with respect to pain or other forms of physical suffering, however. For it is a striking fact that pain and other kinds of physical suffering have a distinctive phenomenology: they *are* immediately recognizable and distinguishable. When we are cold or hungry, tired or in pain, we know that we are in these states, and know in virtue of the distinctive way that these states feel. Indeed, it seems that we can have such knowledge *purely* in virtue of the phenomenal character of the states: we don't need, that

is, additional information about how long we've gone without food or sleep, or about the external temperature, or about the nature of the noxious stimulus, to recognize that we are suffering in these distinctive kinds of ways. More importantly, we don't need to appeal to any evaluative element – any appraisal of the things that constitute the targets of these forms of suffering – in order to distinguish between them, or in order to know that we are cold, tired, hungry, or in pain. Feeling theories of physical suffering thus seem rather more adequate than feeling theories of emotional suffering when it comes to matters of identification and differentiation. So a strong reason for thinking that emotional suffering is evaluative is no reason for thinking that pain and other forms of physical suffering are equally evaluative.

(ii) Metaethics

In the previous section I argued that the values associated with emotions are best understood along rational sentimentalist lines. But the same case cannot be made for the disvalues associated with forms of physical suffering like pain and hunger. This is not because some form of causal or dispositional account of such things is *more* plausible than a rational sentimentalist account. Instead, it is because we don't need to understand such things in terms of *any* response on the subject's behalf. That is: whereas it is very tempting to suppose that disvalues like shamefulness, for example, must have *something* to do with the emotional response of shame, it is not (at all) tempting to suppose that there is any necessary connection between disvalues like bodily damage or lack of food and experiences of pain and hunger. This is because we have more or less objective standards for determining when the body is injured or damaged, or when the body is malnourished. This is why, after all, there can often be puzzling *divergence* between bodily damage and pain – as when physicians determine that someone is injured but doesn't feel pain, or when someone reports pain and yet there are no objective signs of physical injury. Of course, feelings of pain are often very helpful for diagnostic purposes, so that a physician can determine where some injury might be located, or the extent of that injury, as a result of relying upon subjective reports of feelings. But *that* something constitutes an injury or damage – to tissue or nerves, muscles or bones, skin, or major organs – is determined by facts about physiology, normal functioning, and the like, rather than any appeal to the feeling that injuries might generate.

This means that we need not – indeed, should not – understand the disvalues that are associated with physical suffering on rational sentimentalist lines. If so, then we cannot claim that bodily damage necessarily constitutes a reason to feel pain; at least, this claim doesn't follow from metaethical reflection about the concept or property of bodily damage. But then we cannot argue, as we could in the case of emotional suffering, that pain and hunger must be capable of being *rational* responses to their objects. There is no need, in other words, for us to assume that pain and hunger *must* be the kinds of states that can respond to, or perhaps involve, a grasp or recognition of the relevant reason – the kinds of states

than can respond to, or perhaps involve, an awareness of the relevant value, in other words. There is thus no need, from the standpoint of metaethics, to suppose that pain, hunger, coldness, and other forms of physical suffering are responses to, or essentially involve, evaluations.

(iii) *Flexibility*

A third argument that there is an intimate connection between forms of emotional suffering and evaluations, but not between forms of physical suffering and evaluations, rests upon the following striking fact: there is extreme diversity in the objects and events that generate forms of emotional suffering, and an equally extreme diversity in ways of dealing with such things. Pain and other forms of physical suffering, on the other hand, respond to an extremely *limited* class of stimuli and motivate an extremely limited pattern of behavioural responses. I want to now argue that the need for discriminative capacity and behavioural flexibility requires that feelings are bound up with evaluations in emotional suffering; but there is no such need when it comes to pain and physical suffering, and hence no such requirement.

To make this argument, let us remind ourselves that forms of emotional suffering target a particular kind of disvalue. So fear is about danger, grief about loss, shame about social disgrace, guilt about moral fault. But think of the *very many things* of which we might be afraid: physical injury, mental infirmity, burglary, lack of money, nuclear war, right-wing politicians, ruining the dinner party, being sacked, the youths in the park, loss of hair, a mole on the skin, the new boss, the young upstart in the office, the bad review, a Facebook *faux pas*, an infestation of flying ants, the pension fund, the possibility I didn't turn the iron off, whether she'll stand me up on our date. Indeed, these might be the kinds of fears that affect us weekly, or even occupy our daily lives. Now think of the many other negative emotions we can experience on a regular basis as well – such as guilt, shame, anger, disappointment, sadness, regret, remorse, embarrassment, despair – and of the very many targets or objects of all of these forms of suffering too. Take all of these together and you have some idea of the extreme range of emotional suffering and its objects. It is, of course, not surprising that our many negative emotions have very many objects; for our emotions are grounded in things we care about or are concerned about, and typically human beings care about very many things and have a multitude of objects of concern.

Turn now to thinking about the very many ways that we might respond to these objects, for each of these emotions: how we might deal, appropriately, with all of the things that we are afraid of, guilty over, sad about, and so on for all of the other cases of negative emotion. There are, after all, a number of ways in which one might deal with one's guilt – apologies and reparations can take various forms, for instance. It is obvious, moreover, that what is an appropriate response to one kind of danger – such as running away – would be an inappropriate response to other kinds of danger – such as a failure to turn the iron off or a shortfall in the pension fund. So there are myriad objects of emotional suffering,

and myriad ways in which we can respond to these objects. This indicates a clear and obvious need for our emotional systems to have a great deal of *discriminatory capacity* in the detection of all of these different kinds of danger, loss, moral faults, et al., and a very high level of *flexibility* when it comes to ensuring the right kind of behavioural response. This need is best met by putting evaluations at the heart of emotional experience, as I'll now explain.

Some of our emotional responses, such as basic kinds of fear, are responses to things that posed potential dangers to our ancestors – things like looming objects, loud noises, crawling creatures. It is plausible to assume that our fear response to these kinds of things has been determined by natural selection. But through teaching, socialization, and experience we learn that very many other things are associated with danger and learn to be afraid of these things as well. At the same time, we learn about the different ways there are of coping or dealing with such things. In order for such learning to take place, images of these new objects and events, and information about what constitutes an appropriate behavioural response to them, must be linked with feelings of fear in long-term memory, such that future encounters with these things generate feelings of fear and activate the appropriate coping mechanism. But this requires that the subject *represents* such things as dangerous, and for the following two reasons. First, the myriad new dangers we encounter will typically have nothing in the way of physical appearance in common with the old dangers: a daunting exam or a Facebook *faux pas* have nothing visually in common with a rattlesnake or lightning. What they do have in common, of course, is that they all exhibit the formal property of being dangerous. But then the new dangers will only be linked with feelings of fear in long-term memory as a result of some grasp or recognition *that* they are dangerous, and hence similar to the old dangers to this extent. That is, it is only if this point of similarity is recognized that the new dangers will be added to the “calibration file” for danger – namely, the “mental mechanism” that contains a set of representations linked to the value in question, and which can trigger the appropriate behavioural response.¹³

It might be objected that the relevant link could take place *without* evaluation of the object; perhaps we just learn, for instance, that objects of this type generate emotional feelings of that type, and hence the former are linked in long-term memory with the latter, without any need for evaluative mediation. In this way, new dangers are linked to the relevant feelings and responses, but without any need for the subject to recognize the new dangers *as* dangers. However – and this is the second point – we saw earlier that mere feelings are insufficient to identify some feeling as, for instance, a feeling of fear. What we need is an element of intentionality – an assessment of the object of those feelings as dangerous. This means that it is not in fact possible for us to learn to associate these new objects and events with feelings of *fear* without an assessment of the objects and events *as dangerous*. Either way, then, evaluations are a necessary part of associative learning, and hence necessary for us to have discriminative capacity and behavioural flexibility with regard to the very many ways in which things can be dangerous, shameful, morally wrong, and so on.

Compare and contrast this picture with that appropriate to pain and forms of physical suffering like coldness, tiredness, and hunger. These forms of suffering respond to a limited range of inputs. Pain responds to noxious threats to bodily integrity; coldness responds to low temperatures; hunger is triggered by lack of food; tiredness is generated by lack of sleep. And although, as I'll discuss below, there can be very many different kinds of pain, it doesn't seem that there are many ways in which one can be cold, tired, or hungry. There are not, for instance, socially constructed forms of coldness, so that we learn, as a result of socialization and culture, to regard certain things as cold, or for which assessments of coldness require sophisticated cognition. This suggests, at least with respect to coldness, tiredness, and hunger, that there is no need for us to have much in the way of discriminative capacity with respect to the simple disvalues that these things respond to, namely low temperatures, lack of sleep, and lack of food. Moreover, and for a related reason, each motivates a very restricted range of behaviour: seeking warmth in the case of coldness; seeking food in the case of hunger; getting sleep in the case of tiredness. With these forms of physical suffering, there is therefore no need for behavioural flexibility. It's not as if we have myriad ways of dealing with coldness and hunger and tiredness. Instead, these forms of suffering give us simple instructions: Get warm! Get food! Sleep! Given, therefore, a very limited range of stimuli, and a very limited range of behavioural responses, it seems that there is no need for a mechanism that operates to ensure discriminative capacity and behavioural flexibility. There is no need, that is, for coldness, tiredness, and hunger to operate *via* any appraisal of their objects.

Pain is a different case, since there are very many different kinds of pain, with different phenomenal aspects, intensities, and the like. Pain is, in addition, much more susceptible to social and cognitive influences than other forms of physical suffering. As a result, the objects and events that trigger pain are also very varied.¹⁴ However, we can doubt that learning about all of these different objects and events requires anything in the way of evaluation on the subject's part, and in the main because the relevant file in long-term memory that links images of painful objects and events with thoughts or feelings of pain is a file of *causal* rather than *evaluative* links. That is, we can add images of painful objects and events to the file, and thus expand our awareness of all of the various things that are painful, simply in virtue of the fact that such things typically *cause* pain. *That's* what all of these things have in common. We don't, in order to add them to the file, need an additional evaluation of the objects or events as *bad*. For as we also saw earlier, we can identify the relevant feelings of pain *without* needing to employ anything in the way of evaluation or appraisal of the object in question. We know that certain sensations are painful sensations simply by virtue of how they feel, in other words. If so, associative learning need not involve evaluation or assessment of the different things that can bring us pain. And if this is true, then we cannot argue for the view that pains involve evaluations, on the grounds that this is necessary to ensure discriminative capacity and behavioural flexibility.

3

In the previous two sections I have discussed three arguments in support of the view that emotional suffering is essentially evaluative. These are arguments from introspection, metaethics, and flexibility. As we have just seen, none of these arguments supports the idea that pain and forms of physical suffering are evaluative. We can identify and discriminate forms of physical suffering without employing evaluations; we can understand the values to which physical suffering responds without thinking that they have any rational or conceptual link with those responses; and the systems governing pain and other forms of suffering can be suitably discriminative and flexible in the absence of evaluations. As a result, we have good reason to doubt that emotional and physical suffering share the same kind of evaluative structure. If we think that reason-responsiveness requires a grasp or recognition of one's reasons, and that this involves something like an appraisal or evaluation, we thus have a straightforward explanation of our common-sense thought that emotional suffering can be reasonable or unreasonable, whilst physical suffering cannot. But this is not all. The fact that there are these differences between emotional and physical suffering puts, at least to my mind, pressure on a recent and important development in the philosophy of pain, namely, *Evaluativism*.

Evaluativism purports to capture or explain the phenomenology of pain experience by appealing to an evaluation or appraisal of bodily damage or disturbance. As David Bain puts it, "a subject's being in unpleasant pain consists in his (i) undergoing an experience (the pain) that represents a disturbance of a certain sort, and (ii) that same experience additionally representing the disturbance as *bad* for him in a bodily sense."¹⁵ Or, as Brian Cutter and Michael Tye state, "our pain experiences do not just represent the presence of tissue damage, but also (roughly) represent our tissue damage as being bad for us to some degree."¹⁶ On this view, pain would seem to consist of two elements or components. The first, phenomenal state, is usually identified with a *sensation* that represents some kind of bodily disturbance or disorder, malfunction, or damage. The second, evaluative element, is an appraisal of this disturbance, disorder, or damage as bad. And it is this second element that is vital in an explanation of painfulness; as Bain writes, "pain is *unpleasant* ... only because it further represents that bodily disturbance as *bad* for you. If, stepping into [hot bath] water, an asymbolic has a pain that is not unpleasant, that is because, even though the represented disturbance is bad for him, his pain fails to represent it as such; his pain lacks that layer of evaluative content."¹⁷

If the arguments of the previous sections are correct, evaluativism about pain and other forms of physical suffering faces a significant explanatory task. For if emotional and physical suffering are *both* partly constituted by evaluations, as evaluativists suppose, and if the evaluative nature of emotion is central to an explanation of how emotions can be responses to reasons, then why isn't pain equally reason-responsive in the same way? That is, why doesn't the evaluation or appraisal that is supposed to be an inherent part of pain and other forms of physical suffering count

as a kind of grasping of the reasons there are *to be* in pain or to physically suffer, with the result that pain, no less than grief, reflects the subject's take on the reasons that there are for her to be in the relevant state?

It seems to me that evaluativists have two options at this point, neither of which is particularly attractive. (i) They could maintain that pain, coldness, tiredness, hunger, and nausea are indeed rational responses to the relevant dis-values. Or (ii) they could maintain that pain, coldness, tiredness, hunger, and nausea are *not* reason-responsive, despite being partly evaluative states. Option (i) has the considerable cost of rendering evaluativism highly counterintuitive, since common sense and much philosophical tradition holds that pains, itches, shivers, headaches, nausea, and irritation are not rational responses to the relevant stimuli. Option (ii), on the other hand, generates yet another explanatory task for the evaluativist, namely, that of specifying some other factor present in pain and other forms of physical suffering that blocks or undermines the reason-responsiveness of the constituent evaluation, a factor which, moreover, is *not* present or operative in forms of emotional suffering. Neither option, it seems to me, is as plausible as the one that I've argued for throughout the paper: that emotional suffering is reason-responsive, whilst physical suffering is not, because the former, but not the latter, is partly constituted by an evaluation or an appraisal. If so, then the arguments in this paper constitute both an explanation of a common-sense intuition about emotional and physical suffering, and a further argument against evaluativist theories of pain.¹⁸

Notes

- 1 "Feelings that Matter", in *Thinking about Feeling*, Solomon, R. (ed.), Oxford: Oxford University Press (2004), p. 200.
- 2 The fact that some value normatively requires an emotional response doesn't entail, of course, that the response will be forthcoming. All kinds of things can make us fail to feel how we rationally ought to feel.
- 3 Daniel Jacobson writes, in his *Stanford Encyclopedia of Philosophy* entry for "Fitting-Attitude Theories of Value", that "a dispositionalist view of value understands funniness, for instance, in terms of whatever amuses normal humans in standard conditions".
- 4 There are interesting recent discussions as to whether there can be *normative* mistakes in colour perception, generated by the possibility that some of our visual experiences are "cognitively penetrated" by other mental states. For one interesting take on this question, as it relates to emotion, see Jona Vance's paper, "Emotion and the new epistemic challenge from cognitive penetrability", in *Philosophical Studies* 169 (2014), 257–83. Reasons of time and space prevent me from entering into this discussion here, unfortunately.
- 5 This point is nicely made by Daniel Jacobson and Justin D'Arms in "Sentiment and Value", *Ethics* 110 (2000), pp. 725–26.
- 6 He continues: "In recent years, Elizabeth Anderson, Simon Blackburn, Allan Gibbard, Bennett Helm, John McDowell, Kevin Mulligan and David Wiggins (as well as [Daniel] Jacobson and I) have each attempted to explicate at least some evaluative concepts along these lines. Helm (2001), McDowell (1997a, 1997b), Mulligan (1998) and Wiggins (1987) apparently regard the second order sentimentalist schema as applying to evaluative concepts quite generally. Blackburn (1993, 1998) and Anderson (1993) appeal to the more inclusive category of 'attitudes' in most contexts,

- but each treats sentiments as at least one central kind of attitude that can be invoked to explain evaluative judgement. Gibbard (1990) deploys the schema explicitly only with respect to certain examples: shameful, dangerous, and wrong, the last of which receives most of his attention. The schema is second-order because it understands the direct evaluative judgement about X in terms of a normative judgement of 'appropriateness' about our own emotional states, namely, about the having of a sentiment F toward X" ("Two Arguments for Sentimentalism", p. 3).
- 7 A problem initially and forcefully presented by Wlodek Rabinowicz and Toni Rønnow-Rasmussen in "The Strike of the Demon: On Fitting Pro-Attitudes and Value", *Ethics* 114 (2004): 391–423.
 - 8 This condition on justificatory reasons is known as the "internalism requirement" and came to prominence in Bernard Williams's famous paper, "Internal and External Reasons". (Williams, B., *Moral Luck*, Oxford University Press, 1995, 101–13.) In that paper, Williams argued that it is a necessary condition on justificatory reasons for someone to act that it must be possible for that person to act for that reason. Here I propose a similar condition about reasons for emotions.
 - 9 For the foreground-background distinction, see Philip Pettit and Michael Smith's paper, "Backgrounding Desire", *Philosophical Review* 99 (1990), 565–92. In this paper, Pettit and Smith argue that desires can figure in the background of deliberation and action without figuring in the foreground: without, that is, being something that I recognize in my deliberations and decide and act in light of this. Here I want to make a similar point about our being guided by and responsive to reasons. Reason-responsiveness requires that the reason is in the background of one's coming to be in the relevant mental state, without being something that I explicitly recognize in deliberation about whether, e.g., I ought to be in that state for that reason.
 - 10 James, W., "What is an emotion", *Mind* 9 (1884), p. 189. For Lange's view, which he developed independently from James, see his *Ueber Gemuthsbewegungen*, Leipzig: Theodor Thomas (1888).
 - 11 In *Bodily Changes in Pain, Hunger, Fear and Rage*, New York: Appleton (1929), p. 352. See also Errol Bedford's claim, in support of Cannon's point, that the feeling theory is inadequate at the psychological level, as "it presupposes a richness and clarity in the 'inner life' of feeling that it does not possess. What evidence is there for the existence of a multitude of feelings corresponding to the extensive and subtle linguistic differentiation of our vocabulary for discussing emotions?" In "Emotions", *Proceedings of the Aristotelian Society*, 57 (1957), p. 282.
 - 12 We can make the same point using other examples. Consider the phenomenological similarity between the feelings of shame and guilt, or of disappointment and despair, or of grief and sadness. This doesn't mean, of course, that we can't tell these emotions apart: although sometimes it is difficult to identify our emotions, in other cases we are good at knowing when we are feeling ashamed, guilty, disappointed, afraid, grief-struck. But this is because we are good at recognizing the composite state of feeling *plus* evaluation: I know that I'm feeling guilty because the negative affect is directed towards my memory of my morally reprehensible behaviour; I know that this is grief rather than disappointment I am feeling because it is directed towards the loss of someone I loved dearly. I know that this is excitement rather than anxiety because it's directed towards the holiday rather than the examination. Emotions are typically recognizable as a result of the combination of evaluation and feeling rather than by feeling alone.
 - 13 The idea of a "calibration file" as a mechanism in long-term memory that contains representations of the value and an appropriate response is due to Jesse Prinz. See his *Gut Reactions*, Oxford University Press 2004, especially pp. 99–101.
 - 14 Having said this, the appropriate behavioural response to pain is, like the appropriate behavioural response to other forms of physical suffering, a response to a very simple command: make it stop! And there are limited ways of doing this: remove the body

part for the source of the noxious stimuli, or take steps (such as analgesic drugs) to lessen the pain from damage already inflicted.

- 15 "What Makes Pains Unpleasant?", *Philosophical Studies* 166 (2013), p. 82.
- 16 "Tracking Representationalism and the Painfulness of Pain", *Philosophical Issues* 21 (2011), p. 91. See also Bennett Helm, "Felt Evaluations: A Theory of Pleasure and Pain", *American Philosophical Quarterly* (2002), pp. 13–30, for an early presentation of the evaluativist view.
- 17 Ibid.
- 18 Earlier versions of this paper were presented at the University of Edinburgh, at the workshop "Suffering and Cognition" at Paris, Sorbonne, in April 2014, and at the conference on "Suffering and Reason" at the University of Glasgow in 2014. I would like to thank the participants at these events for very helpful feedback. Particular thanks are due to David Bain and Jennifer Corns for stimulating discussion of these issues in Glasgow and very helpful feedback on earlier drafts of the paper. The paper was written while I was Co-Principal Investigator of the *Pain Project* (University of Glasgow), funded by Sam Newlands and Mike Rea's *Problem of Evil in Modern and Contemporary Thought* project (University of Notre Dame), funded by the John Templeton Foundation, and Co-Principal Investigator on the *Value of Suffering Project*, which was also funded by the John Templeton Foundation (Grant ID 44167). I am very grateful to the Foundation for their support for both of these research projects.

5 The placebo effect

Jennifer Corns

1 Introduction

Consider the following case:

Headache: You have a headache. I give you a yellow sugar pill while telling you that it is an aspirin that will make you feel better. You take it and your headache gets better; if asked, you would report a decrease in the unpleasantness of your headache.

This case is taken to be a paradigmatic case of a placebo (the sugar pill) and a placebo effect (the decrease in the unpleasantness of your headache).¹ It is, more particularly, a paradigmatic case of placebo analgesia or placebo for pain—the most well-studied and well-accepted type of placebo effect.

Consider, however, another case:

Disappointment: You are disappointed. I give you a yellow sugar cookie while telling you that it is a well-deserved reward that will make you feel better. You eat it and your disappointment gets better; if asked, you would report a decrease in the unpleasantness of your disappointment.

This is *not* taken to be a case of a placebo (the yellow sugar cookie) or a placebo effect (the decrease in the unpleasantness of your disappointment).

The parallel is troubling for characterizing the placebo effect. Why is an effective sugar pill-cum-kind word a placebo effect for a headache, while an effective sugar cookie-cum-kind word is not a placebo effect for disappointment? In both cases, the sugar-cum-suggestion intervention acts on the same component (unpleasantness, or negative hedonic tone) and results in the same outcome (the decrease of negative hedonic tone). Moreover, in both cases, the sugar-cum-suggestion might well work, in part, because I believe the suggestion: I believe that the cookie will help my disappointment and that the sugar pill will help my headache. Why then is Headache, but not Disappointment, a (paradigmatic) placebo effect?

The problematic parallel is illuminated by contemporary research that has led many to hold that the negative affective component of emotions, like that of pain, is subject to placebo effects. Moreover, this evidence supports the claim

that emotions are subject to placebo effects through the same mechanisms as pain. If that is right, then it looks like Disappointment and Headache are even parallel mechanistically. In contrast, again, with our intuitive judgment that Headache is a placebo effect but Disappointment is not.

In 2005, Petrovic et al. hypothesized that the well-known cases of placebo analgesia were instances of a general placebo effect for the emotions. In particular, they hypothesized (p. 963) that: “placebo analgesia is a special case of reward processing and that placebo treatment could modulate emotional perception in the same way as [it] does pain perception.” The results of their study supported their hypothesis: unpleasantness ratings of non-noxious visual stimuli (pictures) were manipulated by conditioning and verbal suggestion (placebo interventions for emotional unpleasantness) in ways corresponding to the manipulations of unpleasantness ratings of noxious stimuli (placebo interventions for the unpleasantness of pain). The unpleasantness of subjects’ aesthetic experiences was subject to the same effects by suggestion as their pains. Moreover, many of the same brain areas are involved in the modulations. Along with others who cite them, Petrovic et al. (2005) have called effects like this one “emotional placebo.” They conclude their influential study by saying (p. 963): “We suggest that the modulatory processes in placebo are not specific for placebo analgesia, but are rather part of the mechanisms involved in the regulation of emotional processes in general.”

The claim that placebo analgesia is a type of emotional placebo appears to be the emerging consensus. In discussing Petrovic et al. (2005) and further studies, Benedetti and Amanzio (2013) write (p. 416) that the results taken together suggest “functional-anatomical relationships between placebo analgesia and emotional regulation in which top-down modulation of the pain or emotional network is implemented.” Many experiences, emotional experiences in particular, are unpleasant, and that unpleasantness appears to be subject to placebo effects. The unpleasantness of emotions, it is increasingly accepted, is as subject to placebo treatments as the unpleasantness of pain—and through the same mechanisms.

Nonetheless: Headache is taken to be a placebo effect and Disappointment is not. This parallel is problematic enough for our everyday intuitions about placebo effects, but consider the following case:

Aspirin: You have a headache. I give you a yellow aspirin while telling you that it is an aspirin that will make you feel better. You take it and your headache gets better; if asked, you would report a decrease in the unpleasantness of your headache.

Aspirin is a paradigmatic case of analgesia. Unlike Disappointment, we might think, Aspirin is unproblematically compared to Headache. The parallel between Disappointment and Headache is troubling because the effects of a sugar pill on the negative hedonic tone of a pain and the effects of a sugar cookie on the negative hedonic tone of an emotion seem relevantly similar—but we consider

one, and not the other, a placebo effect. Characterizing the effects of a sugar pill on negative hedonic tone as a placebo effect while denying that the effects of an *aspirin* on negative hedonic tone are a placebo effect, however, does not seem problematic. These appear to be relevantly different. As technology has advanced, however, the parallel between Aspirin and Headache has also become problematic.

It has turned out that many of the same pathways, neurochemicals, and brain areas involved in paradigmatic cases of analgesia, as in Aspirin, are likewise involved in paradigmatic cases of placebo analgesia, as in Headache. Thus, Benedetti (2009) notes (p. 255) “one of the most practical implications of recent neurobiological advances in placebo research is the possibility to induce, at least in some circumstances, drug-like effects without the administration of drugs.” And these similarities, as reviewed in Benedetti’s important book, are not just limited to the ultimate effects on symptoms, but the mechanistic means by which symptom changes come about. When you take an aspirin and an aspirin-shaped sugar pill, much the same physiological things, relevant to the decrease of your unpleasantness, occur.

If we accept these empirical results, the parallel between Headache and Aspirin is as problematic for characterizing the placebo effect as that between Headache and Disappointment. The parallel between Headache and Disappointment may not be problematic for some placebo researchers, but it is problematic for our everyday intuitions. On the other hand, the parallel between Aspirin and Headache may not be problematic for our everyday intuitions about placebo effects, but it is problematic for placebo effect researchers.

These problematic parallels raise the following question: if two interventions result in the same effect, why is one a placebo effect while the other is not? Philosophers have, as yet, made scant contributions to grappling with this question. Inquiries into placebo treatments across multiple disciplines are increasing, and all, either implicitly or explicitly, require that there is some satisfactory answer. Despite these problematic parallels, it is assumed by these researchers that the placebo effect can be usefully identified and—as far as ethically permissible—exploited for treatment.

In this chapter, I argue that the assumption that there is a distinct class of placebo effects² that is usefully identified for inquiry and treatment is unfounded. In the following two sections, I discuss what the placebo effect is supposed to be: in Section 2, I present cases offered as placebo effects in the literature, before critically evaluating characterizations of the placebo effect meant to capture them in Section 3. In short, I argue that characterizations of the placebo effect are subject to insurmountable scope problems. Moreover, identifying the placebo effect is discussed in the literature as being useful for three main purposes: (i) randomized controlled trials; (ii) mechanistic inquiries; and (iii) treatment. In Section 4, I argue that a distinct class of placebo effects is not useful for any of these three purposes. Not only do we lack a successful characterization of the placebo effect, we do not need one. In Section 5, I offer four suggestions for why, despite the presented arguments, it may remain intuitive that some token or type of effect is

a placebo effect; why, despite the arguments in Sections 3 and 4, we might still find it intuitive that Headache is a placebo effect, while neither Disappointment nor Aspirin are. Though I lack the space to fully resolve them, it is my hope that our lingering intuitions about placebo effects can ultimately be explained away along my suggested lines. As I conclude in Section 6, it is my hope that having explained away these intuitions, we will be better able to embrace all effective treatments as equally legitimate. More effective treatment for pain, among other maladies, may then be in the offing.

2 Cases

To consider whether placebo effects are a usefully identified class of effects, it is appropriate to first consider supposed examples of such effects. This section presents 11 types of effects considered to be placebo effects, before critically evaluating attempted characterizations of the placebo effect as such in Section 3.³

Here is a list of 11 types of effects offered as placebo effects in the literature:

2.1 *Placebo analgesia*

Already mentioned in the introduction, placebo analgesia is the most well-known, well-accepted, and extensively documented type of placebo effect. Here are three of many findings about placebo analgesia:

- a Somewhat unintuitively (to me, anyway), placebos tend to be more effective for severe pain than for mild pain (Hoffman et al. 2005).
- b Placebo pills were more effective at reducing pain intensity when subjects were told that they were \$2.50 per pill as against “discounted” pills for \$0.10. Though note that *both* were effective: 61% of those taking the discounted pills reported pain reduction compared to 85% of the others (Waber et al. 2008).
- c Conscious conditioning, by pairing distinct visual stimulations with distinct intensities of noxious stimulations, facilitated the effectiveness of subsequent non-conscious placebo interventions for pain (Jensen et al. 2012). Thus, placebo analgesia appears to be sometimes non-conscious—albeit subsequent to conscious conditioning.

2.2 *Placebo for movement*

Verbal suggestion alone (e.g., Pollo et al. 2002), and both verbal suggestion and conditioning (e.g., Benedetti et al. 2003), have been effective for Parkinson’s patients’ movement. Pollo et al. (2008) also found increased muscle work (“objective” measure) and decreased muscle fatigue reported (“subjective” measure) by conditioning and suggestion. Note that in this study, the subjective measure required conditioning, while the objective measure was manipulable by suggestion alone.

2.3 Placebo for cough

Eccles (2006) found that only 15% of the reduction in cough from taking cough medicine was attributable to the “active pharmacological component”; they conclude that 75% of the effectiveness of cough syrup is placebo.

2.4 Placebo and chemotherapy

Stockhurst et al. (2000) found anticipatory nausea and vomiting in chemotherapy patients, i.e., patients who had previously undergone chemotherapy treatment began to feel nauseated and to vomit on subsequent occasions when chemotherapy drugs were expected but not yet administered. Anticipatory food aversion in chemotherapy patients has also been found (Jacobsen et al. 1993).

2.5 Placebo surgery

- a “Sham” arthroscopic surgery for osteoarthritis of the knee is as effective as “real” surgery (Moseley et al. 2002). “Real” surgery is a common treatment for the condition, involving ligation (tying up an artery) or debridement (removing dead tissue), whereas the “sham” surgery involved a few small incisions to the skin on the knee.
- b “Sham” surgery (without ligation) for angina pectoris (chest pain) has been found to be as, and perhaps more, successful than “real” surgery (with ligation). See Moerman (2002) for references and extensive discussion.

2.6 Placebo allergies

Oral challenge involves testing those with known allergies to new substances and doses of their known allergens. Both “subjective” (e.g., dizziness) and “objective” (e.g., skin rash and lesions) symptoms have been found in response to placebo oral challenge (Liccardi et al. 2004, Lombardi et al. 2008). Note that the pattern of placebo allergy symptoms often does not match the patient’s original adverse drug reactions.

2.7 Placebo anti-depressants

Placebo anti-depressant pills have been found to mirror changes in brain activations caused by anti-depressant pills when tested at both 1 and 6 weeks. Moreover, the types of brain changes found mirror the types of depression treatment against which the placebo was tested. See Benedetti et al. (2005) for discussion.

2.8 Placebo for inflammatory and immune-related diseases

A number of inflammatory and immune-related diseases are subject to placebo effects. These include: asthma, cancer, Crohn’s disease, chronic fatigue syndrome, duodenal ulcer, irritable bowel syndrome, multiple sclerosis, and ulcerative

colitis. For a review, references, and the methodologies of the supporting studies, see Pacheco-Lopez et al. (2006).

2.9 Reverse placebo effects

- a In Levine et al. (2006), participants sat in a motion-sickness inducing device. Both reported nausea (“subjective” measure) and gastric dysrhythmia (an “objective” measure of nausea) were *increased* by a placebo pill and the suggestion that it was effective for motion sickness—as compared to both controls and those told the pill would make their motion sickness *worse*.
- b In Howard and Hughes (2008), participants told that they were going to inhale a scent that would *facilitate* relaxation underwent an *increase* in galvanic skin response, whereas those told that the inhaled scent would *inhibit* relaxation underwent a *decrease*. Decrease in galvanic skin response was used as an objective measure of relaxation. Note that there was no significant difference between groups on the “subjective” measure: self-reported relaxation.

2.10 Placebo for itch

Intensity and unpleasantness of itch has been found to be subject to placebo suggestion—indeed *more* so, in this study, than pain (van Laarhoven et al. 2011).

2.11 Non-medical placebos

Placebo coffee and placebo alcohol have been found to cause effects mimicking those of coffee and alcohol. These and effects like them are often called “non-medical” placebo effects. See Stewart-Williams and Podd (2004) for references and discussion.

Though these and many more examples of the placebo effect are individually identified throughout the literature, there is no agreed upon way of characterizing the effect of which they are all supposed to be examples.

3 Characterizations

Characterizations of the placebo effect assume that there is some difference between placebo effects and non-placebo effects.⁴ Placebo effects and non-placebo effects are presumed to be mutually exclusive. Even if we allow some problem cases at the borders, a successful characterization must therefore be one that is broad enough to get (at least) the paradigmatic placebo effects in, but narrow enough to leave all the paradigmatically non-placebo effects out. Problems including the (at least) paradigmatic Xs and excluding the (at least) paradigmatic non-Xs in a characterization of X are problems of scope.

Scope problems for characterizations of the placebo effects are legion. Some problems of scope concern particular cases. As in the introduction, for instance, Headache, Disappointment, and Aspirin seem parallel in ways that challenge

our different intuitions about these cases. Similarly, consider the cases of Placebo Movement (2) in the previous section. How are these cases different from good results seen from hiring a personal trainer—a presumably paradigmatically non-placebo effect? If my personal trainer whispers in my ear, “You can do it! You’re stronger than you think!”, then is my improved lifting ability or reported decrease in muscle fatigue a placebo effect? If these effects of suggestion offered to me in a gym by a trainer are not placebo effects, why do placebo effects include the effects of suggestion offered in a laboratory, by a researcher, to a person with Parkinson’s? A successful characterization of the placebo effect must provide an answer.

In addition to the many particular problematic cases like these, there are some general scope problems for characterizations of the placebo effect. I briefly consider two: bias effects and psychological effects.

The practical problem of separating the bias effects from placebo effects is well-recognized by placebo researchers. Biases are, roughly, systematic errors that result in incorrect assessments. Hrobjartsson et al. (2011) identify nine types of bias that they think present a challenge to the reliability of placebo effect findings, including response bias in particular. Response bias is the skewing of answers as a result of a participant’s desire to give a “correct” or desired answer. Benedetti (2009) similarly offers a table of “factors that can cause the false impression of a placebo effect,” and these include both observer and patient (reporting) bias. Note that the assumption underlying these discussions is that, though bias effects and placebo effects are difficult to distinguish in practice, they are different in principle.

The problem is not only one in practice, however, but also in theory: it is hard to see how one can characterize placebo effects such that bias effects are excluded. The problem is especially pronounced for placebo analgesia, since reporting is the gold standard for pain. We do not appear able to distinguish a report bias here at all, even in theory.⁵ But the problem runs deeper than reporting and extends wider than placebo analgesia.

The difficulty, in short, is that some outcomes are mediated by psychological states and considered placebo effects, while other outcomes mediated by psychological states are instead considered bias effects—but there seems no principled way of distinguishing the placeboogenic psychological mediators from the others. Recall that expectation is the dominant explanatory mechanism currently on offer for placebo effects. What we need, then, is a non-arbitrary characterization of placebo effects that includes the effects of *expectation* of an effective treatment, but excludes the effects of *desire* for an effective treatment as a bias effect. But why do expectation-driven effects count as placebo effects, while desire-driven effects count as bias effects? Similar concerns arise for other biases mediated by participants’ psychological states.⁶

This difficulty is related to the problem of scope arising for characterizing placebo effects as distinct from psychological effects. The core concern is that all psychological effects will be classed as placebo effects. As Stewart-Williams and Podd (2004) note, characterizations that emphasize the physical inertness as distinct from the psychological effectiveness of placebo interventions seem

to render all psychological effects as placebo effects. They drily note that this “is clearly an undesirable conclusion” (325). The difficulty is to avoid this problem in a non-arbitrary way. The problem of distinguishing placebo effects from psychological effects in general is particularly pressing if we want to be able to characterize placebo effects *within* psychology; if all psychological effects are placebo effects, then this will be impossible.⁷

These scope problems are taken by placebo researchers *as* problems to be solved—and the common approach to solving them is an appeal to some further distinction. This further distinction is applied to one of three aspects of placebo treatments: the intervention (i.e., the placebo, e.g., a saline solution), the outcome (i.e., the placebo effect, e.g., a decrease in reported unpleasantness), or the process from the intervention to the outcome (i.e., the placebo healing process, e.g., the decrease in unpleasantness as resulting from the injected saline solution). Different placebo researchers often emphasize these different aspects of the placebo treatment when attempting characterizations of placebo effects. These different aspects are clearly interrelated, and characterizations of any one should be consistent with, and are often dependent on, the others. The distinction concerning an aspect of placebo treatments is offered that is, directly or in turn, deployed to characterize placebo effects. To successfully facilitate characterization of the placebo effect, however, any deployed distinction must not only be tenable in its own right, but the two distinguished features must both not be true of the placebo effect.⁸

Below is a list of six distinctions influentially deployed in the literature to characterize the placebo effect by applying it to some aspect of placebo treatment.⁹ Insurmountable problems, I argue, arise for all. Many of these problems have been noted by others. Here are the six (the aspect of placebo treatment for which the distinction purportedly holds is emphasized in parentheses):

- 1 Active as distinct from inactive/inert (intervention)
- 2 Real/genuine as distinct from fake/illusory/sham (intervention)
- 3 Pharmacological as distinct from psychosocial (intervention)
- 4 Specific as distinct from non-specific (intervention and outcome)
- 5 Treatment effect as distinct from psychological or context effect (outcome)
- 6 Legitimate healing as distinct from illegitimate healing (process).

All of these distinctions are problematic when deployed to characterize the placebo effect. I consider each in turn.

***Regarding 1: Active as distinct from inactive/inert
(intervention emphasis)***

Deployment: Placebo interventions are characterized as inactive or inert treatments as distinct from active treatments. To characterize placebo effects, it is further claimed that placebo effects are the effects resulting from placebo interventions.

The active/inert distinction is the one most traditionally deployed to characterize the placebo effect. Deploying this distinction to characterize the placebo effect faces the most well-known problem for characterizations of the placebo effect in the literature: the placebo paradox. Koshi and Short (2007) nicely put the problem this way (10):

Does such a thing as placebo (inactive substance or treatment) exist? ... There is a paradox with our definitions of placebo and placebo effect. If placebos are inert substances, they cannot cause an effect. If an effect occurs, the placebos are not inert.

The distinction between active and inactive/inert is mutually exclusive, and we cannot deploy this distinction to characterize placebos (and in turn placebo effects) without running into a version of the placebo paradox. Placebos turn out to be both inert/inactive (using the distinction) and active (as causes of the placebo effect), but these are mutually exclusive. Something has to give.

The most common strategy in the literature in responding to the Placebo Paradox is to continue to characterize placebo interventions by deploying the distinction between active and inert, but to resist characterizing the placebo effect as the effects of the placebo intervention. On these views, the placebo effect is not the effect of the placebo (an inert substance) but something else, e.g., an effect of the context in which the placebo intervention (inert substance) is given (as in deployed distinction 4 discussed below).¹⁰ The benefit of this strategy is that inert things stay inert and no paradox arises. A seemingly unnoticed problematic consequence of this strategy is that the placebo effect is no longer the effect of the placebo: a rather odd strain on the language.

Most important for present purposes, however, is that if the Placebo Paradox pushes us to limit our deployment of the distinction between active and inert to interventions, then it is of no further interest. Granting that placebos are inert treatments will not yield a successful characterization of placebo effects. Even if we grant the deployment of this distinction for characterizing placebo interventions—which it is not clear that we should do—we are still left without a characterization of placebo effects.

Regarding 2: Real/genuine as distinct from fake/illusory/sham (intervention emphasis)

Deployment: Placebo interventions are characterized as fake/illusory/sham interventions as distinct from real interventions. To characterize placebo effects, it is further claimed that placebo effects are the effects resulting from placebo interventions.

Though less well-noted, the very same problem arises for the deployment of this distinction as that arising for the deployment of the distinction between active and inactive: the placebo intervention is characterized as fake or unreal, but it must be real insofar as it is the cause of the placebo effect. If we assume that no effects can have unreal causes, then deploying this distinction to characterize

the placebo effect lands us in paradox. One may again respond to this problem by liberating the placebo effect from placebo interventions, with the same results: i) the odd consequence that the placebo effect is then not the effect of the placebo, and ii) we remain without a characterization of the placebo effect.

Regarding 3: Pharmacological as distinct from psychosocial (intervention emphasis)

Deployment: Placebo interventions are characterized as psychosocial interventions as distinct from pharmacological interventions. To characterize placebo effects, it is further claimed that placebo effects are the effects resulting from placebo interventions.

The problem with this increasingly popular characterization is, in short, that not only are pharmacological and psychosocial interventions not mutually exclusive, but most (if not all) pharmacological interventions are also psychosocial interventions. If most (perhaps all) pharmacological interventions are psychosocial, how can we successfully deploy this distinction to characterize placebos or, in turn, placebo effects?

Consider the paradigmatically non-placebo intervention of aspirin in Aspirin. Aspirin is both a pharmacological and psychosocial treatment. It is a pharmacological treatment in virtue of its chemical structure—it is in virtue of its chemical structure that we consider it to be effective for headaches. But it is also a psychosocial treatment: I have taken those little yellow pills many times in the past to good effect, starting with when my mother gave them to me as a child. When aspirin is effective, it is partly in virtue of these psychosocial features—this history of aspirin-taking makes aspirin-taking more effective. Aspirin is thus both a pharmacological and psychosocial intervention. A similar story can be told for (at least) most pharmacological interventions.¹¹

Here is one response: the contributions of the pharmacological and psychosocial are separable. This response requires both that there are distinct pharmacological and psychosocial aspects of a treatment and that the distinct features constituting these aspects make discrete contributions to the outcome. The placebo is the psychosocial treatment aspect, and the placebo effect is the effect resulting from this aspect alone.

There is, however, no reason to think that pharmacological and psychosocial treatment contributions are discrete, and some reason to think that they are not. As noted in the introduction, paradigmatic placebo interventions are now recognized to often involve much the same explanatory mechanistic activity as their paradigmatically non-placebo, pharmacological counterparts: pharmacological and psychosocial interventions often utilize the same pathways, neurochemicals, and neural areas. It is therefore implausible to think that supposedly distinct treatment aspects of a treatment like Aspirin make distinct contributions. As is true for the contribution, so is true for the effect. Our best research suggests that the effects of my taking an aspirin are the *combined* result of the pharmacological and psychosocial aspects of the intervention.¹²

Notice, importantly, that those treatments administered in the hidden condition in an open-hidden paradigm are no exception here. Open-hidden paradigms involve comparing the results of administering a drug openly—with the full awareness of the patient—and surreptitiously—without the patient being aware of the administration. This second condition is called the hidden condition. It is sometimes claimed that the difference in effect between open and hidden administrations is the difference between the pharmacological treatment and the pharmacological-cum-psychosocial treatment. The hidden administration is taken to simply subtract all psychological aspects, the so-called “psychological contamination” of the treatment.

Despite its many other virtues, however, the open-hidden paradigm will not salvage the distinction between pharmacological and psychosocial for characterizing the placebo effect. “Hidden” interventions are still administered to a particular patient, in a particular room, at a particular time; i.e., all treatments are given in a particular psychosocial situation to a patient who has a particular psychological profile. Hidden administrations do not reduce the patient’s psychology to null. The difference between open and hidden administration is therefore *not* the purely subtracted difference between the pharmacological and the psychosocial. It is, rather, a comparison of one pharmacological-cum-psychosocial intervention with another. Some aspects of the psychosocial situation are different—in particular, whether the subject or patient knows that they are receiving an intervention at the precise moment they receive it—but by no means have all aspects been eliminated; memories and personal history, for instance, remain equally relevant. The same is true for chemicals snuck into patients’ food, given to infants, or to patients in a coma.

For those unconvinced, however, that the pharmacological and psychosocial effects cannot be neatly separated, notice that a more direct and perhaps less contentious problem for the deployment of this distinction remains. If all psychosocial interventions are placebo interventions, and all effects of psychosocial interventions are placebo effects, then far too many psychological effects are going to be placebo effects. That is: any psychological effect resulting from a psychosocial intervention would be a placebo effect. The same problem arises if we require a pharmacological intervention for something to be a placebo effect: there is no way to distinguish psychological effects that are placebo and those that are not. Effective psychology turns out to be effective placebo. Thus the deployment of this distinction appears unable to solve one of the original scope problems for which it is needed.

Regarding 4: Specific as distinct from non-specific (intervention or outcome)

Two deployments:

- a Placebo interventions are characterized as non-specific interventions. To characterize placebo effects, it is further claimed that placebo effects are the effects resulting from placebo interventions (intervention emphasis).

- b Placebo effects are characterized as non-specific effects as distinct from specific effects (outcome emphasis).

A version of the placebo paradox also arises for the deployment of the distinction between specific and non-specific. When used to characterize placebos (intervention): if the intervention is non-specific, then how can it also be the specific cause of the placebo effect? When used to characterize the placebo effect (outcome): if the effect is non-specific, then how can it also be specified as the placebo effect? Indeed, as discussed in the following section, the study of the placebo effect is largely driven by (and thought useful because of) identification(s) of the mechanism(s) of the placebo effects, which assumes not only that there are specific effects, but the in-principle specification of the mechanisms of these effects.

One might attempt to respond by differently defining “specific.” One popular such alternative definition of “specific” and “non-specific” is offered by Bootzin and Bailey (2005). Specific effects are explicitly defined as those effects predicted by “the theory,” whereas non-specific effects are defined as those *not* predicted by “the theory.”

But this type of reliance on theory-relativity is problematic on other grounds.¹³ The resultant characterization of the placebo effect is as an effect that one’s theory does not predict. This is surely unacceptable. If a placebo effect is simply an identified effect that your theory does not adequately account for, the right response is to change your theory—not to call the unaccounted-for effect a “placebo effect.” The appropriate response to unpredicted outcomes is revision, not special labels for the outcomes. As we should have learned since the inception of gate-control theory, for instance, if our theory of pain cannot give a substantive causal role to cognition, then our theory of pain needs to change. It is inappropriate to simply lump and label all unaccounted-for effects; instead, we need to revise our theories to account for them.

Regarding 5: Treatment effect vs context effect (outcome)

Deployment: Placebo effects are characterized as psychological or context effects as distinct from treatment effects.

The first problem with the deployment of this distinction is that all treatment effects are given in a context that contributes to the effects of the treatment. Just as the psychological cannot be neatly subtracted from the pharmacological (deployed distinction 3), the effects of the treatment cannot be neatly subtracted from the effects of the context in which the treatment is given.

A further and more significant problem, however, is that it is not clear how we are supposed even to make the distinction between treatment effects and context effects that is being deployed. We might attempt to isolate the effects of and only of identified contextual features as we attempted to do for the psychosocial features of a treatment. But specifying the contextual features of an intervention as against the, say, *substance* of an intervention is arbitrary insofar as both are granted to have effects.

We might think that seemingly superficial properties of the treatment are merely contextual and not substantive if anything is; e.g., the color or shape of a drug intervention. Notice, however, that such seemingly superficial features are sometimes directly posited as treatments and tested against a placebo. So, for instance, Howard and Hughes (2008) used tea tree oil as a “placebo smell” for lavender oil. We might have thought that smell was a contextual feature of a treatment if anything is, but smell sometimes is the substantive treatment (in this case, for anxiety). The moral is that whether color or shape is a substantive or merely contextual feature of the treatment depends on one’s theory of how the treatment works.

As with deploying distinction 4: if it turns out that some feature that we thought was merely superficial substantially contributes to an outcome, then our theory of the outcome should respect that contribution. Insofar as our theory does not predict, explain, or allow for the discovered contribution, we should change our theory. Again: the appropriate response to unpredicted outcomes is revision, not special labels for the outcomes. If a contextual effect as against a treatment effect—and, in turn, a placebo effect as against a non-placebo effect—is just an effect that our theory does not account for, then the right response is to change our theory to provide an account.

Regarding 6: Legitimate healing vs illegitimate healing (process emphasis)

Deployment: Placebo effects are those effects resulting from an illegitimate healing process (from intervention to outcome), as distinct from legitimate healing processes (from intervention to outcome).

If deploying this distinction results in a successful characterization of placebo effects as the result of an effective type of healing, then there seems no reason remaining not to consider placebo healing processes as legitimate. Efficacy should be granted as sufficient for legitimacy. But if placebo healing processes are legitimate types of healing, then they cannot be successfully characterized by deploying this distinction. If it is granted that effective healing is legitimate healing in virtue of its being effective, then the paradoxical problem of distinctness for this deployment is straightforward. And, again, it is hard to see why we should resist considering effective healing to be legitimate healing.

Those who deploy this distinction in fact appear to be attempting to legitimate placebo treatments by characterizing them as types of healing that society wrongly thinks are illegitimate. Thus Kaptchuk (2011) emphasizes placebo treatments as types of “illegitimate” healing—but this is an almost honorary term for Kaptchuk. Illegitimate healing for Kaptchuk (2011) appears to be effective, ritualized healing often not acknowledged as legitimate by those societies dominated exclusively by Western medicine. The idea that placebo effects should be characterized simply as those that the current medical establishment is biased against is also, however, not something that Kaptchuk wants to embrace. There is a difference between placebo treatments and quackery—even for Kaptchuk.

Once it is admitted that placebo healing processes *are* legitimate, however, then it should be admitted that this distinction is not successfully deployed to characterize the placebo effect.

In summary: characterizing the placebo effect is subject to recognized scope problems that require the deployment of a further distinction to solve, but deployments of the representative and influential six considered all flounder. It may be that deploying some further distinction will nonetheless succeed, but I think this is unlikely. Similar problems plague all those of which I am aware. I think it is instead time to consider why we should expect some future characterization to succeed.

One reason we might expect a future characterization to succeed is that we think, as is often claimed, that our current identifications of placebo effects as against non-placebo effects are useful. Is it useful to identify any token or type of effect as a placebo effect?

I want to grant that utility would be a good reason to think that some as yet undeveloped characterization can succeed where the extant fail. Moreover, utility would not only render the search for a successful characterization of placebo effects as such worthwhile, but it would also remain worthwhile to continue to identify particular instances of the placebo effect while the search remains incomplete. (Consider that lacking a successful characterization of, say, chairs, does not make it any less appropriate for me to ask you to bring me one when I want to sit down.) In the following section, I argue against the utility of identifying an effect or class of effects as placebo effects.

4 Utility

The utility of identifying some effects as placebo effects may initially seem supported by the increasing number of multidisciplinary inquiries into such effects. So, Benedetti (2008) writes (p. 34):

The widespread use of the word placebo ["placebo"] in the medical literature, and its use [the use of placebos] in many experimental procedures, is clear evidence of the importance of this phenomenon in modern biomedical science. Knowledge of the placebo effect is therefore essential in modern medicine.

The word "placebo" does indeed undergo widespread use in the medical literature, but widespread usage of a word does not entail the utility of distinguishing its supposed referent. The word "witch" was widespread for quite some time—categorizing groups of women as witches was nonetheless not a useful practice, and inquiries aimed at gaining knowledge *about witches* should not have been encouraged. Usage of a term may *prima facie* support utility, but it does not entail it. We need to look at the way in which the term is employed.

There are three main purposes in the literature for which identifying a distinct placebo effect, or class of placebo effects, is offered as useful. These are (i) the

supposedly crucial role of the placebo effect in randomized controlled trials; (ii) inquiries into the supposed mechanisms of the placebo effect; and (iii) the supposed treatment benefits of exploiting the placebo effect. I argue that identifying a placebo effect, or some class of placebo effects, is lacking in utility for all three of these purposes. From its failure to be useful for the purposes offered for it by its advocates, I will conclude that identifying a distinctive class of placebo effects is not useful.¹⁴

Consider first and at most length the supposedly crucial role of the placebo effect in randomized, controlled trials (RCTs). RCTs are considered the gold standard for clinical trials, i.e., the gold standard for evaluating new treatments. In an RCT, persons in a population of interest (e.g., persons with asthma) are randomized to one of at least two interventions (the treatment “arms”), and the outcomes resulting from these distinct interventions are compared. RCTs can vary (among other ways) in the number of interventions compared, the types of factors used to select the population of interest, the types of controls used in distinguishing interventions, and the type of analysis used to evaluate the outcomes.

Most important for present purposes is that RCTs almost always involve a “placebo arm”: some of the population of interest is randomly assigned to a supposed placebo intervention.¹⁵ The outcome for those receiving the placebo intervention is taken to be the placebo effect, and it is taken to be a sign of legitimacy for a treatment that the outcome(s) in its arm(s) is greater than that measured in the placebo arm(s). Any treatment intervention that is not more effective than a placebo intervention is considered, indeed, to be ineffective.

The first problem for the supposedly crucial role of identifying the placebo effect as such for RCTs is that, as exemplified in Section 2, placebo effects are *effects*. Greene et al. (2008), nicely puts the problem this way (p. 94):

The construct underlying the use of placebos in clinical trials is the assumption that a supposed active drug or treatment has to be at least superior to a placebo in producing positive outcomes. The implicit assumption of this research design is that placebos have little, if any, efficacy ... However, recent research has shown that placebos do, in fact, elicit definite biological as well as behavioral responses from patients within a wide variety of medical conditions. ... placebos almost never have a zero effect.

If placebo effects are effects of treatment, then attempting to subtract them from treatment effects is a mug’s game. But the role of placebo effects in RCTs assumes that we can do just that—and indeed, that doing so is crucial for evaluating the legitimacy of treatment interventions. The problem is not that we shouldn’t compare treatment arms; it is that the effects in the so-called placebo arms are not zero.

One attempted response is to argue that, despite the seeming evidence to the contrary, placebo effects *do* have a zero effect—or, anyway, statistically significantly zero. This claim is what has made Hrobjartsson and Gotzsche’s (2001, 2004, 2010) meta-analyses so infamous. Hrobjartsson and Gotzsche (2001,

2004, 2010) characterize placebo effects as any effects in the “placebo arm” of a treatment as suggested above, and they characterize the “placebo arm” as whatever is so-called within the studies included in their meta-analyses. They then consider such effects across multiple studies, spanning multiple conditions. They do find some clinically significant placebo effects in particular conditions—in particular, for pain—but even these are massively heterogeneous.¹⁶ Considered across all conditions, moreover, they find that there is no clinically significant placebo effect. If the placebo effect is no effect at all, then perhaps it can be compared to the treatment effect as needed for RCTs after all.

These meta-analyses have been subject to many additional critiques from placebo researchers.¹⁷ Most focus on whether Hrobjartsson and Gotzsche have been studying the “real” placebo effect. Benedetti (2009), in a typical example, writes (p. 6):

by taking the recent neurobiological discoveries about real psychological placebo phenomena into consideration, the scientific community is becoming more and more aware that it is necessary to differentiate real placebo effects from other phenomena that are present in clinical trials.

The objection, in short, is that lumping all measured outcomes in the “placebo arm” conflates the “real” placebo effect with other distinct effects: e.g., regression to the mean, habituation, co-interventions, and more. Hrobjartsson and Gotzsche themselves seem, at times, to recognize this, writing (2010, pp. 14–15): “It is a question of definition whether our review evaluates the ‘placebo effect.’ ... We do not exclude the possibility of important effects of other aspects in the patient-provider interaction.” But what is the placebo effect as against the effects of other aspects of the patient-provider interaction?

It becomes clear that the lack of a successful characterization of the placebo effect argued for in Section 3 is important for their supposedly crucial role in RCTs: if the placebo effect cannot be successfully characterized, then it cannot be compared to the effects of the intervention being tested. As argued in the previous section, there is no extant successful distinction between placebo effects and treatment effects. Of course, one could compare any one treatment arm with any other treatment arm one wishes—but that will not show that we have determined that a treatment is effective by comparing its efficacy with the placebo effect. What is needed for this latter claim is that the outcome in at least one of the arms is, indeed, *the placebo effect*; since it is the utility of identifying the placebo effect *as such* in RCTs that is being considered. It is hard to believe that RCTs involve the successful comparison of the placebo effect with treatment effects if they cannot even be successfully distinguished from one another.

One response is to appeal to the broad characterization of the placebo effect (not discussed in Section 3) seemingly implicit in Hrobjartsson and Gotzsche’s work as *whatever* outcomes we find in the placebo arm. But appealing to the “placebo arm” will only help if we have some characterization of the placebo intervention that can in turn be used to characterize the placebo effect. Without

a successful characterization of the placebo that can be used to characterize the placebo effects, the broad label remains empty. It is in this way that Hrobjartsson and Gotzsche have left themselves open to the charge that they have lumped many effects together under one empty label by those placebo researchers who do identify some particular class of effects as placebo effects.

Indeed, without a successful characterization of the placebo effect, it is hard to see how we can evaluate whether Hrobjartsson and Gotzsche's work is telling us anything relevant about it. If Hrobjartsson and Gotzsche lump together many effects manifest in the outcomes of the "placebo arms," then their work cannot be employed to salvage the utility of identifying a distinct class of placebo effects useful for RCTs. If the meta-analyses are not about the placebo effect in particular, as other placebo researchers have argued they are not, then their work is irrelevant for present purposes. If their work does not identify or analyze a distinct effect at all—the "real" placebo effect or otherwise—then it is even less relevant. I conclude that these meta-analyses will not support the crucial role of identifying the placebo effect for RCTs.

Nonetheless, evaluating new treatments obviously remains crucial: I suggest that we re-conceptualize RCTs without placebo arms or effects, but instead as involving the comparison of outcomes of interventions that are similar and different in specified ways. This we can do without distinguishing a distinct placebo effect or class of such effects. It is good methodology to compare the outcomes of similar treatments; it is valuable to compare aspirin pills with sugar pills, for example, and to compare arthroscopic surgery with and without debridement. Such comparisons will tell us many things about how different interventions work and inform decisions about which interventions are safe, effective, and worth developing. All this we can, and should, do without identifying a class of placebo effects. Nunn (2009) nicely puts the point this way (p. 338):

In a post-placebo era, experiments will simply compare something with something else. That is, they will compare experimental conditions: one group gets these conditions and another group gets those conditions. The report of every methodologically acceptable experiment will describe the conditions that have been compared ... There will be no more hiding behind the skirts of the emperor's new placebo. ... Eventually we will have standard descriptions for commonly compared things. Legislation will reflect those standards. We gain transparency, honesty, and clarity.

Turn second to the purported utility of identifying placebo effects for inquiries into their supposed mechanisms. It is now nigh-universally held among placebo researchers that there is not one mechanism implicated in placebo effects, but many. The question is then whether it is useful to investigate these many mechanisms *as* placebo-effect mechanisms.

Expectation and conditioning receive the most attention and are the most commonly offered explanatory mechanisms of the placebo effect: is it useful to investigate expectation and conditioning *as* mechanisms of the placebo effect?

One problem with so doing is that it is not at all clear that we should think of expectation and conditioning as distinct mechanisms of distinct subtypes of placebo effects. As forcefully argued by both Stewart-Williams and Podd (2004) and Greene et al. (2008), we do better not to think of these as mutually exclusive. If these arguments are sound, they present problems for the view that expectation and conditioning are alternative mechanistic explanations of placebo effects. We would, I suggest, do better to directly investigate conditioning, expectations, and their relationship, as against pitting them against each other in explaining supposed placebo effects.

Let us focus particularly on expectation as an explanatory mechanism of the placebo effect. Participants in RCTs are usually simply assumed to have expectations of receiving effective treatment, and these expectations are in turn taken by many to be the explanatory mechanism for measured outcomes in the “placebo arms.” In a qualitative study, however, Kaptchuk et al. (2009) found that almost all participants in the target RCT explicitly denied that they had any expectation of treatment benefit. They did, however, report states like hope, concern, and increased attention to their symptoms. Both clinical and laboratory experiments on the placebo effect often simply assume that participants form expectations when given verbal suggestions, but whether they have in fact formed any positive expectations is rarely tested. This calls into question the role of expectations as explanatory of supposed placebo effects. Insofar as subjects do not have expectations—but instead perhaps have a range of distinct attitudes like hope or concern—using results from placebo studies that simply assume expectation to be causally explanatory in the found effects will contaminate our models and understanding of expectation.

Expectations may well play an important role in the efficacy of an intervention, but studies that investigate expectation as a mechanism for placebo effects as such do not appear helpful for better deciphering that role. We would, I suggest, do better to investigate expectation and the role it may play in treatment outcomes without modeling expectation as a placebo-effect mechanism.

I suggest that what goes for inquiries into expectation, the most commonly offered explanatory mechanism of the placebo effect, goes for inquiries into the other mechanisms implicated in supposed placebo effects. Though space precludes consideration of each in turn, *I have yet to encounter a posited mechanism that would not be at least as usefully investigated if liberated from the notion of a distinct class of placebo effects*. If there is no such distinct class of effects, this is to be expected. Moreover, we can continue to investigate expectation, conditioning, somatic attention, and any other mechanisms implicated in the healing process without construing any of these as placebo-effect mechanisms.

Third and finally, turn to the purported treatment benefits of exploiting the supposed placebo effect. Improved treatment is often mentioned as one important benefit of identifying, and better understanding, placebo effects. Studying placebo effects *as such* is therefore taken to be useful for treatment purposes. One often reads that improving our understanding of the placebo effect will help us sharpen existing tools and identify new ones. What exactly these tools are, however, or what the placebo effect as such has to do with them is rarely made clear.

As an example, consider Miller et al. (2009). They conclude (p. 532): “We suggest that it is fruitful from a theoretical perspective to consider the placebo effect, in the context of interpersonal healing, as a central tool in the art of medicine.” But what exactly is this supposed tool? Miller et al. (2009) recognize that interpersonal healing involves many factors. One particularly important factor in interpersonal healing is the patient-healer relationship. And the patient-healer relationship itself involves multiple factors. Let us grant that these many factors of the patient-healer relationship are important: what does that have to do with a distinct class of placebo effects? Identifying a distinct class of placebo effects is not needed to further exploit potentially beneficial features of the patient-healer relationship.

Going further, notions of the placebo effects, along with placebo interventions, seem to hinder the development and maximal use of currently underused tools. Consider: what is to be gained from considering the patient-healer relationship a placebo and the salutary effects of this relationship as a placebo effect? Suggesting that a treatment is a placebo suggests that it is *mere* placebo. Speaking positively about treatments, for instance, may well be beneficial for healing. How is it useful, however, to identify the result of speaking positively about treatments as a placebo effect or label positive-speaking as a (mere) placebo? If creating ceremony is beneficial for healing, as another for instance, then it ought to be encouraged. How is it encouraging to classify ceremonies as placebos or the results of ceremonies as placebo effects? Identifying a treatment tool as a placebo, or the results of its use a placebo effect seems not to promote, but instead to disparage further use.

Treatment will be better improved, I suggest, by ridding ourselves of the idea that there is a distinct class of interventions (placebos) and outcomes (placebo effects) and, instead, exploiting effective treatments and using our best tools *full stop*. A beneficial treatment tool is better exhorted by *not* identifying it as a placebo. Far from being useful for improving treatment, the idea that there is a distinct class of placebo effects impedes our embrace of novel and effective treatments.

In summary, not only do we lack a successful characterization of the placebo effect, but identifying supposed placebo effects, or a class of effects, as such is not useful for RCTs, mechanistic inquiries, or treatment.¹⁸ Some future characterization of the placebo effect may succeed where these have failed, and there may be some utility in identifying some effect, or class of effects, as yet unnoticed. There seems no reason, however, to think so. Extant characterizations of the placebo effect fail, and the purposes for which the placebo effect is supposedly crucial would be better accomplished without identifying any effects, or class of effects, as placebo effects.

5 Intuitions

I have argued that there is no extant successful characterization of placebo effects and no known utility from distinguishing a class of such effects. Despite

these arguments, it will likely remain intuitive that some effects are placebo effects while others are not. Our intuitions about, e.g., Headache, Aspirin, and Disappointment will likely linger. Given that there is no successful characterization of the placebo effect, and that no token or type of effect is usefully identified as a placebo effect as such, our lingering intuitions that there *are* such effects cry out for explanation. This done, we can, hopefully, consider these intuitions to be explained *away*.

The intuitions that need explaining, however, are unlikely to be uniform. In particular, the intuitions of the placebo researcher may often differ from the non-expert. I have hitherto focused only on expert characterizations of the placebo effects and those effects taken by placebo researchers to be placebo effects. It is unlikely, however, that expert and non-expert intuitions are uniform. It is, for instance, unlikely that the non-expert will think that *all* the cases presented in Section 2 are placebo effects—despite the fact that they are offered in the literature as such by placebo researchers. Moreover, placebo researchers' intuitions about particular cases are likely based on explicit theories, and these explicit theories are likely to differ across researchers. We should expect, then, that intuitions about which effects are placebo effects will differ both between and among experts and non-experts.

Despite these differences, in this section I now suggest four groups of background beliefs that I suspect contribute, if variously, to both expert and non-expert intuitions about placebo effects.¹⁹

(1) Dualism: I suggest that beliefs according to which some effects are the result of mind-body interactions, while others are the result of body-body interactions, sometimes contribute to the intuition that some token or type of mind-body interaction effect is a placebo effect.

The smell of dualism pervades both non-expert talk of the placebo effect and many of the claims, characterizations, and paradigms offered by the placebo researcher. There is, for example, sometimes talk within placebo research of a “central commander,” standing over and above the physiological body, whose heeded commands are explanatory of the placebo effects (e.g., Benedetti 2009, p. 218, and Pollo et al. 2008, p. 385). In describing placebo effects, the non-expert describes placebo effects as, for example, “mind over matter” or “all in your head.” Both types of talk lend themselves to dualistic interpretation: the commander who is above, but not part of, the physical body, and the mind that is over, but not itself, matter. Dualistic beliefs about mind-body interactions like these are, I suggest, often a contributor to both expert and non-expert intuitions that an effect (or class of effects) is a placebo effect (or type of placebo effect).

(2) Bottom-up biases: I suggest that beliefs according to which some processes are bottom up while others are top down sometimes contribute to the intuition that some token or type of top-down processing effect is a placebo effect.

That some effects are placebo effects seems to me to be intuitive for many placebo researchers because they find it so *unintuitive* that some processes are subject to top-down modulation. It is commonly held that some things, like

lifting my arm or my face turning red, can be caused and altered from the “top down”; i.e., by my intentions, thoughts, or emotions. But other things, like feeling pain, coughing, or being drunk, it might be thought, can only be caused or altered from the “bottom up.” Such ideas are most likely to be held by placebo researchers who have explicit theories about which processes are “bottom up” and which are “top down.” When effects that a researcher expects to be the result of a bottom-up process appear, instead, to be caused or modulated from the top down, the resultant effect may then be considered to be a placebo effect. Ideas concerning top-down and bottom-up processing are specialized and part of an explicit theory, and such notions are therefore unlikely to contribute to the intuitions of the non-expert. I suggest, however, that effects which are theorized by experts to be the results of bottom-up processing are more likely to be considered placebo effects when they are found, instead, to result from processes theorized to be top down.

(3) Indirect control: I suggest that beliefs according to which some outcomes can be indirectly controlled by beliefs (or other mental states), while others cannot, sometimes contribute to the intuition that some token or type of indirectly controlled effect is a placebo effect.

Background beliefs about beliefs play a large role in non-expert intuitions about the placebo effect. Consider Headache: the non-expert is likely to think that it is your belief about the sugar pill that explains the placebo effect: the sugar pill works because you believe it is an aspirin. Indeed, the non-expert is likely to believe that the outcome is a placebo effect precisely *because* it is caused by a belief in treatment efficacy. One might construe the role of beliefs about efficacy as a kind of indirect control over the treatment outcomes. Insofar as the outcome is explained by one’s beliefs, it is in that sense under one’s control. Insofar as one cannot simply choose to believe that some intervention is an effective treatment, however, one’s control over the outcome is in that sense indirect. Ideas about the role of beliefs in indirectly controlling symptoms are often, I think, major contributors to non-expert intuitions that an effect is a placebo effect. The non-expert is much more likely to think that an effect is a placebo effect if they also think that the effect is (causally) explained by a belief.

Going further, one might think that a characterization of the placebo effect that tracks non-expert intuitions can be built out of this suggestion of indirect control. One might, that is, think the non-expert definition of the placebo effect is (or is something much like) the following:

Y is a placebo effect iff it is caused by some X (the placebo) at least in part because a subject believes that X will cause Y.²⁰

Note first that this is straightforwardly inadequate as a characterization of the placebo effect as understood by the expert for three main reasons. First, as in the case of placebo movement listed in Section 2, there is sometimes no relevant X in placebo studies; in this case, there is suggestion and the surreptitious switching of weights, but no intervention recognized as an intervention by the subjects.

Second, some placebo effects remain even when participants are explicitly told, and presumably believe, that an intervention is a placebo; i.e., told that it is an intervention that is not effective for causing Y. Third and related are RCTs. In RCTs, persons are told they may have been randomly assigned to a placebo treatment and they may or may not believe they are receiving a placebo—but outcomes in the placebo arm are all classed as placebo effects independent of the mental states of participants.

While the non-expert, by contrast, may simply deny that effects in these three types of cases are placebo effects, there are nonetheless problems with the above as an adequate characterization of even the non-expert notion. First, the non-expert is quite willing to talk about the importance of beliefs and desires for the efficacy of some treatment interventions that they nonetheless do not consider placebos and whose effects they do not class as placebo effects. For instance, many non-experts accept the importance of the belief that one can win the fight against cancer for the efficacy of chemotherapy; but I daresay that no one thereby believes that when chemotherapy is successful the result is a placebo effect. So the above definition does not seem to identify a sufficient condition that tracks non-expert intuitions. Moreover, it is not clear that it identifies a necessary condition. As discussed below, qualitative studies (e.g., Bishop et al. 2012) seem to suggest that beliefs are not central to the non-expert notion. Instead, the non-expert notion seems to include treatment effects as placebo effects whenever the intervention is something the non-expert considers to be illegitimate, fake, or unreal—independent of their beliefs about the intervention. For example: the non-expert may believe that acupuncture will not be effective for their back pain. But if an undergone acupuncture treatment is nonetheless effective, the non-expert will *still* be likely to consider the effect a placebo effect if it turns out that the acupuncture treatment they received was only “sham” acupuncture (acupuncture using non-penetrating, retractable needles). In this way, the non-expert notion appears to be akin to the expert judgment when evaluating RCTs: it is the type of intervention that matters, not the beliefs of the subject receiving it.

In short, while I want to agree that if a belief plays a role in causing an effect the non-expert will be more likely to consider that effect to be a placebo effect, I do not think this fully characterizes the non-expert notion, and it is hopeless as a characterization of the expert notion. For these reasons, the suggestion here is the modest one that if a non-expert thinks that an effect is indirectly controlled by a person’s beliefs (or desires, expectations, and so on), then that will sometimes contribute to their intuition that an effect is a placebo effect.

(4) Legitimacy: I suggest that beliefs according to which some interventions are legitimate causes of treatment outcomes, while others are illegitimate, sometimes contribute to the intuition that some token or type of illegitimately caused effect is a placebo effect.

Background beliefs about which interventions are legitimate are, I think, key contributors to placebo-effect intuitions. Beliefs about legitimacy are evident in the placebo research discussed above, perhaps most strikingly in the supposedly

crucial role of the placebo effect in RCTs. The role of beliefs about legitimacy is perhaps even more prominent in non-expert intuitions. In a qualitative study of participants in a placebo acupuncture trial, Bishop et al. (2012) conclude (p. 4) that “[p]lacebo acupuncture as a scientific necessity was imbued with undesirable characteristics through the language used to describe it, as a hoax, a phony, and a fake.” They go on to write (p. 6): “Our own findings highlight the need to examine dominant discourses of placebos and suggest that societal level interventions may be needed to challenge what appear to be established notions that placebos are fake treatments that have illusory effects.” Background beliefs about legitimacy are, I suggest, crucial to our intuitions that some effects are placebo effects.

Consider the contribution of background beliefs about legitimacy to our intuitions about Headache, Disappointment, and Aspirin. A cookie and a kind word is a legitimate intervention for disappointment—not a placebo effect. An aspirin is a legitimate intervention for a headache—not a placebo effect either. A sugar pill and a suggestion for a headache is, however, illegitimate—any effect caused by this intervention is a placebo effect.

Notice further that 1-3 may all be contributing reasons for a person’s beliefs about illegitimacy. An expert who believes that a process which her theory predicts is “bottom up” is instead being manipulated from the “top down” may take that very belief as a reason to think the effect of the modulation is, therefore, less legitimate. A non-expert who believes that some particular treatment only works because they believe it works may conclude that the treatment is therefore illegitimate. The belief that an effect is “all in your head” can, for the expert and non-expert alike, undermine legitimacy. And these beliefs about illegitimacy, in turn, I am suggesting, are major contributors to the explanation of our intuitions about which effects are placebo effects.

In summary, I suggest that all four mentioned groups of background beliefs contribute—individually and combined, variously for experts and non-experts—to intuitions about placebo effects. Background beliefs about legitimacy, I suggest, are particularly explanatory. These suggestions are truly suggestions. Nonetheless, I think they are suggestive of the sort of background beliefs that explain our intuition that some effects, or class of effects, are placebo effects. Given that there is no successful characterization of the placebo effect and no utility to be gained by identifying any effects as such, explaining away any intuitions to the contrary is the final task.

6 Conclusion

In closing, note that ridding ourselves of our intuitions about placebo effects—or at least ridding our medical practice from classifying certain effects as placebo effects—may have practical benefits. Separate is (haven’t we learned?) rarely equal. Insofar as we distinguish placebo treatments from other treatments, our biases against their legitimacy, despite their efficacy, are likely to persist. One important benefit of ceasing to class any effects as placebo effects is that we will then be freer to embrace all effective treatments as equally legitimate.

Such freedom promises benefits both for researchers investigating and developing new treatments and for patients seeking them—for pain and beyond.²¹

Notes

- 1 Note that some placebo researchers, e.g. Price et al. (2008), distinguish between the effects of a placebo intervention on an individual and a group and call the former a “placebo response” and only the latter a “placebo effect.” For purposes here, I will not distinguish between these and will use only the more common “placebo effect.”
- 2 Note that I will move back and forth between talk of the placebo effect and a class of placebo effects. The arguments in text, if successful, refute both the idea that there is a single placebo effect and a more plausible placebo-effect pluralism, according to which there is a class of placebo effects.
- 3 These examples are not exhaustive, but I take them to represent a fair range.
- 4 Note that in this section I restrict my attention to those characterizations of the placebo effect offered in the placebo literature. Though I think similar or the same problems for any non-expert characterizations will arise, non-expert characterizations of the placebo effect are explicitly addressed only in Section 5.
- 5 This is not skepticism about reporting bias; only skepticism that this type of bias can be distinguished from placebo effects—especially but not only in the placebo analgesia case.
- 6 It is perhaps for this reason that Kaptchuk et al. (2009) list response bias as a recognized conceptual mechanism of the placebo effect! The attempt to distinguish bias from genuine effect is simply given up. Since bias is a measuring error, such a move by a prominent placebo researcher—however idiosyncratic—is by itself perhaps an indication that the identification of a class of placebo effects is problematic.
- 7 Similar scope problems arise for other areas of medicine, e.g., psycho-neuro-endocrino-immunology.
- 8 To successfully characterize the placebo effect by distinguishing between some features X and Y, X and Y cannot both be true of the placebo effect.
- 9 This list is not exhaustive, but each of these six is either dominant or influential, and taken together present what I take to be a representative range. It is also worth noting that Moerman (2002) identifies and characterizes a related phenomenon, the meaning response, which is influential in placebo effect research. Moerman’s meaning response is the response by a patient to the context in which healing occurs (p. 16). Moerman is explicit that he does not take himself to be characterizing the placebo effect, though he does think that his meaning response (i) explains many of the effects that have been classed as placebo effects by others and (ii) is a distinct type of effect.
- 10 An appeal to a belief in an inert substance that is causally responsible for the effect is considered under distinction 3 (as proffered by the expert) and in Section 5 (as intuited by the folk).
- 11 One might also hold a reductionist view according to which all psychosocial interventions are also pharmacological, but this more controversial claim is not needed for present purposes.
- 12 Again, one might further hold the more controversial claim that all of psychosocial contributions are reducible to pharmacological contributions. If such reductionism is accepted, the distinction will be even less successful as deployed for present purposes.
- 13 Compare to Alfano and Miller (forthcoming), who seem to think that *all* characterizations of the placebo are problematically theory-relative, perhaps even necessarily so. While I think this is a worry for discussed distinctions 4 and 5, (i) the others discussed do not seem to me problematically theory-relative, and (ii) I am not convinced that theory-relativity *as such* is problematic.

- 14 This may not settle the issue once and for all: someone may later identify some utility in identifying a class of placebo effects. But showing that the class is not useful for the extant purposes offered in the literature will, I hope, be enough to settle the issue for now.
- 15 Turner (2012) suggests that we think of these as PCTs: placebo-controlled trials. Indeed, Turner thinks that we should retain the notion of placebo comparisons despite the problems he rightly recognizes for the utility of notions of placebo and placebo effects. Though space precludes further discussion, Turner gives no convincing reason for thinking the notion of placebo comparison has any utility that cannot instead be secured by clearly specifying the controls used in any particular control group. Talk of placebo comparison, as Turner recognizes for “placebo” and “placebo effect,” obscures without benefit.
- 16 Hrobjartsson and Gotzsche are tempted to think that these and the other patient-reported outcomes for which they find a heterogeneous placebo effect are not placebo effects at all but reporting bias. As in Section 3, it is hard to see how we can successfully characterize placebo effects to exclude reporting bias—especially for placebo analgesia.
- 17 E.g., Pacheco-Lopez et al. (2006) and Price et al. (2008).
- 18 As discussed in text, this is not to deny the utility of comparing outcomes across treatment arms, identifying and investigating mechanisms implicated in treatment efficacy like expectation, or better exploiting underused treatment tools like positive-speaking.
- 19 Moerman (2002) argues (especially pp. 134–139) that the key to understanding our intuitions about so-called placebo effects is the background idea that people are machines. Moerman holds that we are surprised by those effects that cannot be accommodated by the idea that we are machines, and it is these surprising effects that we identify as placebo effects. Moreover, he argues, the existence of these effects shows that we are wrong to think that we are machines. Contra Moerman, however, I do not think that the idea that we are machines explains our intuitions about placebo effects; no causal relationships are, by themselves, problematic for mechanistic ideas. Though, as in text, the type of machine a person thinks that humans are may matter for their intuitions about placebo effects.
- 20 I am grateful to David Bain for offering me a near-identical formulation and for helpful discussion of this point.
- 21 I am thankful to David Bain, Fabrizio Benedetti, Robert Cowan, Colin Klein, and Daniel Moerman for helpful comments on earlier drafts of this paper; to participants of the “Suffering and Cognition” workshop at Maison de la Recherche, the Consciousness and Self-Consciousness Research Seminar Series at the University of Warwick, and the Seminar Series at the University of York for their helpful feedback on presented parts of this paper; and to the John Templeton Foundation and the Value of Suffering Project, with PIs David Bain and Michael Brady, for funding this research.

6 What is the affective component of pain?

Jesse Prinz

1 Introduction

It is a truism that people and other living things usually try to avoid pain. But why? The standard answer from folk psychology is that pain is unpleasant. We avoid it because it feels bad. This folk view has a counterpart in the leading scientific theories of pain. According to the prevailing view, pain has two components: sensory and affective. The affective component is posited to explain why we avoid pain. Under some conditions, these components can be dissociated, and a sensory experience of pain without an affective component is said to be non-aversive. The affective component is what makes pain bad. But what is the affective component? That simple and fundamental question has received surprisingly little attention.

Here I will consider some proposals about the nature of the affective component. The word “affect” is often used as a scientific stand-in for emotion, and it is tempting to say that the affective component of pain is an emotion. But which emotion? Is it an emotion distinctive to pain or a more basic emotion that also arises in other contexts? If so, which one? The most plausible answers to these questions will be considered, and found to be wanting, in Section 2. I will then propose that the affective component is not an emotion, but rather what is sometimes called “negative valence.” There are different theories of valence, and I will defend a particular approach and argue that it provides a theory of the affective component of pain. The advantages of this theory will be reviewed in Section 3. I also will address a pressing objection (Section 4): if the valence theory proves to be right, it may follow that pain is not unpleasant, contrary to folk intuitions about pain and pain avoidance. This may sound absurd, but I will argue that there are theoretical advantages to this implication.

2 Identifying the affective component

2.1 *The sensory/affective theory of pain*

The idea that pain has a component structure has been in the literature for almost half a century. It was influentially advanced by Melzack and his collaborators in the 1960s (e.g., Melzack and Casey, 1968), though they found seeds of the

idea in Charles Sherrington's work at the turn of the 20th century. In this early work, Melzack and Casey distinguish what they call sensory-discriminative components and motivational-affective components. These are now usually called sensory and affective components. The sensory components refer to the fact that pain characteristically carries information about specific kinds of changes in the body and their locations. For example, someone might feel a sharp, shooting sensation on the arch of her right foot. On its own, the sensory component does not account for the badness of pain. For that, Melzack and Casey postulate an affective component. They do not offer an explicit definition there, but refer to pain's "unpleasant, affective, quality."

As evidence for these two components, Melzack and Casey point to examples of dissociations. In particular, they identify cases where the sensory qualities of pain seem to be intact, but averseness is lost. One example is soldiers who sustain extensive injuries but then insist that they don't need any medication. Melzack and Casey interpret these as cases of people who can sense their injuries but have no negative affect because they are overjoyed to have survived a battle. They also cite evidence from lobotomy patients, who have normal discrimination thresholds but don't seem to mind painful stimuli. The same, they say, is true of some people with congenital insensitivity to pain: they correctly identify the sharpness of a pinprick but don't withdraw. One could say of these individuals that they experience pain without one of its normal components, or one could say that the missing component is necessary for calling it pain. This is essentially a verbal issue, and I don't take a strong side. The key substantive point is that pain ordinarily has a negativity in addition to its sensory qualities. Calling this negativity a "component" means that it is ordinarily bound to the sensory component, and thus, the two function as a kind of unit. They can be considered parts of a complex mental state. Melzack and Casey helped bring attention to pain's complexity.

The componential view is now widely accepted. Melzack (1983; 1987) went on to develop an influential pain measurement tool – the McGill Pain Questionnaire – which carefully distinguishes sensory and affective dimensions. Subsequent work has added further support to the conclusion that these can function independently. For example, some analgesics, such as morphine, reduce pain affect without reducing pain sensation (Kupers et al., 1991).

From the beginning, Melzack postulated different brain pathways for processing sensory and affective information in pain. That work has been advanced by the advent of modern functional imaging techniques, such as PET (positron emission tomography) and fMRI (functional magnetic resonance imaging). Modern imaging demonstrates that different aspects of pain correspond to different anatomical lesions. The sensory dimensions of pain are associated with somatosensory areas and other brain structures known to contribute to bodily perception. For example, PET scans have revealed that hypnotic reduction of the sensory aspects of pain, but not affective aspects, modulate activity in the somatosensory cortex (Hofbauer et al., 2001). The somatosensory cortex, along with the anterior insula, were also associated with sensory aspects in a PET study (Peyton et al., 1999); the study modulated attention during heat-induced pain

and found that these regions remain active regardless of the attention condition, suggesting that they serve to detect noxious stimuli even when we are distracted from pain's unpleasantness.

In contrast, the affective dimension has most frequently been associated with activity in anterior cingulate cortex. For example, anterior cingulate is modulated by hypnotic reduction of negative affect in pain (Rainville et al., 1997). In another PET study, Tölle et al. (1999) confirmed that portions of the anterior cingulate correlate with the unpleasantness of heat-induced pain. They found that some other areas of the cingulate, as well as areas of the frontal inferior cortex and the thalamus, are associated with sensory detection of pain. Sawamoto et al. (2000) also associated heat-induced pain unpleasantness with anterior cingulate, and they found activity in posterior insula as well. Fulbright et al. (2001) used cold-induced pain in an fMRI study and confirmed that anterior cingulate correlates with pain negativity, along with right superior frontal gyrus and the right cuneus. Younger et al. (2011) found that brain volumes in anterior cingulate reduce along with pain unpleasantness when morphine is given as an analgesic for chronic back pain. Anterior cingulate lesions are also known to reduce the affective component of pain without disrupting detection of painful stimuli (Hurt and Ballantine, 1974), and surgical removal of the cingulate relieves pain unpleasantness (Foltz and White, 1962).

Despite subtle differences in these findings, the basic pattern is the same and is unsurprising. Sensory components of pain are most frequently associated with responses in the brain's somatosensory system, and affective components are most frequently associated with anterior cingulate. The anterior cingulate has been implicated in various aspects of human psychology. These include conflict monitoring (Botvinick et al., 1999), error detection (Kiehl et al., 2000), and especially emotion (e.g., Damasio and van Hoesen, 1983; Devinsky et al., 1995; Etkin et al., 2011; Lindquist et al., 2012). The link between cingulate and emotion fits well with the idea that the cingulate underwrites the affective component of pain. It points to the suggestion that the affective component is a kind of emotional state. I turn to that attractive hypothesis now.

2.2 Is the affective component an emotion?

There is a long history linking pain and emotion. Within philosophy, pain has often been regarded as a basic building block of emotions. This view has been defended by Aristotle, Spinoza, and Hume, who each claim that every emotion is a form of pleasure or pain. Within this tradition, pain and pleasure can be characterized as the most basic emotions, upon which other emotions are constructed. For example, Spinoza (1677/1992: Part III) defines pity as "pain accompanied by the idea of ill that has happened to another whom we think of as like ourselves," and he defines shame as "pain accompanied by the idea of some action of ours that we think others censure." When talking about emotions, Spinoza often uses the Latin term "affectus," and the corresponding English term, "affect," has been used as a synonym for "emotion" in many contexts ever since.

When Melzack and Casey (1968) invoke an affective dimension of pain, they may have had emotions in mind. They resist the term “emotion,” which had been disfavored by behaviorists, and they emphasize motivation – a point to which I will return. In the years since, it has become commonplace to treat the affective component of pain as an emotional state. For example, in a *Science* article reviewing work on the affective component, we find Price (2000: 1769) saying, “Part of the affective dimension of pain is the moment-by-moment unpleasantness of pain, made up of emotional feelings.” Likewise, the International Association for the Study of Pain (IASP, 1979) defines pain as, “An unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage.”

It is interesting to note that this emotional interpretation of the sensory-affective divide effectively reverses the Aristotelian idea that emotions have pain as a component. They imply, instead, that pain has an emotional component. This curious inversion deserves reflection, and I will return to it as well. But first I want to consider what an emotional theory of the affective component might look like.

If one equates pain’s affective component with emotion, the first question that arises is, which one? Which emotion could be a component part of pain? Does pain comprise a proprietary emotion, or does it rather comprise an emotion that also arises in other contexts, such as one of those found on familiar lists of basic emotions? This issue has received surprisingly little attention, so it is difficult to explicate endorsements of any specific answer. Many authors talk vaguely about pain having an emotional component without commenting on the identity of that component. Others mention pain’s interactions with specific emotions but remain silent on the question of whether any of these emotions are components of pain. Here I will try to give voice to several proposals that invite evaluation.

Let’s begin with the idea that there is a proprietary emotion involved in pain. In the literature, this conjecture is rarely stated explicitly, but it is strongly suggested by a group of researchers who investigate the way pain is expressed facially. Facial expressions have a distinguished history in emotion research. Distinctive expressions are widely regarded as indicating basic emotions (Darwin, 1872; Ekman, 1993). Thus, finding a distinctive pain expression might be taken as evidence that the emotional component of pain is unique to pain. This is precisely what a number of researchers have sought to establish (LeResche, 1982; Craig and Patrick, 1985; Prkachin, 1992; Williams, 2003). Drawing on such research, Craig (2003) has even gone so far as to argue that pain itself qualifies as a basic emotion.

Research on the expression of pain usually uses Ekman and Friesen’s (1978) Facial Action Coding System (FACS), a technique for observationally measuring the various muscular changes that make up a facial expression. By analyzing photographs or film clips, researchers have found that there are particular “facial actions” that indicate with statistically significant frequency when people are in pain. This has been done with adults and infants, with naturally occurring pain or laboratory-controlled pain, and using a range of pain inducers. It has also been

suggested that the facial actions associated with pain differ, when considered collectively, from those of other emotions that have been studied using FACS. In a review of the literature, Williams (2002: table 1) distinguishes the expression of pain from anger, sadness, and fear. Such findings encourage the idea that pain encompasses a distinctive emotion.

We should be cautious, however, in inferring a distinctive emotion of pain from this research on facial expressions. Studies purporting to show a pain expression often analyze images of people in pain without direct comparison to other emotions. This makes it difficult to establish whether pain has a distinctive expression. Other studies simply assume that pain has a different expression from other emotions, based on previous literature, without showing that observers can correctly distinguish pain (e.g., LeResche and Dworkin, 1988). When participants have to distinguish pain from other emotions, results have been mixed. For example, in a study where a picture of a pain expression was compared to the prototypical expressions of six basic emotions, pain was correctly recognized with a hit rate of 58.8%, taking error into account. This hit rate is appreciably lower than the hit rate for the six basic emotions (mean = 75.5%). In this study, a forced-choice method was used, increasing the probability of success. A somewhat higher hit rate for pain expressions (74.2%) was obtained by Simon et al. (2008). Their stimuli, however, were posed by actors and carefully vetted for discriminability. Moreover, they base their hit rate on the fact that participants gave the posed pain expressions a higher intensity rating on a pain scale than on scales for other emotions. This does not entail that they identified the faces as pain, and, indeed, the mean intensity rating was at the middle of the pain scale, which could imply a lack of confidence. In another forced choice study, researchers compared expressions of pain to extreme sexual pleasure (Hughes and Nicholson, 2008). Pain pictures were correctly identified 75.3% of the time, which is middling, given that chance performance is 50%. Other work suggests that people perform at chance when trying to distinguish genuine pain expressions from faked expressions (Bartlett et al., 2014). This failure is significant because evolved facial expressions are presumed to be difficult to fake (Frank, 1988). In sum, success at pain recognition may depend on the presence of carefully selected and highly contrasting alternatives; studies have not shown that there is a naturally occurring pain expression that can be consistently distinguished from mock pain, sexual pleasure, or a variety of other emotions.

Proponents of the view that pain has a distinctive facial expression might dismiss these doubts and remind us that pain is associated with facial actions that distinguish it from other emotions. Research using the FACS system is supposed to bear that out. This research, however, has been far from decisive. In her review of the literature, Williams (2002) looked at 16 facial actions observed in 15 studies. Only three of these actions were observed in more than half of the studies. The first is a raising of the cheeks. This facial action is not distinctive of pain. In fact, it is a well-documented correlate of happiness (Ekman et al., 1990). This may sound paradoxical, but it turns out that people tend to briefly smile when in pain (Kunz et al., 2009). This may be a way to regulate pain or to communicate

to one's self and others when a pain is not a cause for grave concern. There is also research suggesting that this facial action correlates better with the sensory aspect of pain than the affective aspect (Kunz et al., 2012).

The other two facial actions do correlate with the affective aspect of pain. One of these, which is defined disjunctively as wrinkling the nose or raising the lip, is also a standard component of the prototypical disgust expression. The other, a lowering of the brow, is also not unique to pain; it has been found to correlate with a range of negative emotions, including disgust, sadness, anger, and fear (Gosselin et al., 1995: 90). So these two facial actions are hardly sufficient for establishing that pain involves a distinctive emotion.

To this, one might reply that the concurrence of three actions – cheek raising, nose wrinkling, and brow lowering – is distinctive of pain. But the literature does not firmly demonstrate that these concur as part of a single expression; they might unfold in time. The presence of three facial actions could indicate an alternation or even a blend of two emotions. Thus, while suggestive, extant literature on facial expressions does not establish a pattern robust enough to confirm the conjecture that pain contains a proprietary emotion.

Moving to the opposite extreme, one might propose that the affective component of pain is not a single proprietary emotion, but rather an interchangeable range of more familiar emotions. The fact that pain is expressed using facial actions associated with multiple negative emotions is consistent with this possibility. In the literature, it is again difficult to find an explicit defense of this hypothesis, but various authors come close. Mounce et al. (2010) conducted a principal components analysis of scales associated with the affective dimension of pain. The factor that accounted for the greatest proportion of the variance (12.5%) was termed “general distress.” This might seem like a distinctive emotion in its own right, but the authors define it disjunctively as “anxiety, depression, negative affectivity, and fear of negative evaluations from others” (p. 714). Similar findings are reported by Vancleef et al. (2009). Consistent with this, Fernandez and Turk (1992) report a study in which the therapeutic reduction of chronic pain correlated with reductions in three emotions: anger, fear, and sadness.

The observation that pain correlates with several different emotions is important, but not decisive, in establishing that pain has variable emotional components. A competing hypothesis would state that these emotions are causally related to pains rather than constituent parts (see Craig, 1999). Having a painful experience can be irritating, worrying, and depressing, and alleviation of pain could reduce these side-effects. Methodologically, it is challenging to figure out how to distinguish this causal interpretation from the constituency hypothesis. One potential method would be to look at the time-course of these emotions. For example, sadness seems most likely when pain endures for a certain amount of time (see Fishbain et al., 1997, on depression and chronic pain). Another method would be to see whether similar stimuli cause different emotions in different contexts. For example, anger might be most likely when pain can be blamed on another (see Scott et al., 2013, on anger, injustice, and pain). Such findings,

which are amply supported by anecdotal experience, favor the view that these emotions are context-sensitive sequelae of pain rather than constituents.

I am inclined to conclude that anger and sadness are causally related to pain, but not constituent parts. When it comes to fear, however, a parallel verdict may seem hasty. Of all the classic basic emotions, fear is most plausibly postulated to be the affective component of pain. Fear is an emotion designed to deal with dangers and threats. Pain does indicate threatening ailments or injuries, and therefore it wouldn't be surprising to see fear and pain closely linked. There are also empirical reasons for taking this hypothesis seriously. Fear induction (or, more accurately, anticipatory anxiety) is associated with lower pain thresholds (Rhudy and Meager, 2000). Conversely, the affective component of pain may diminish in contexts where people learn that a pain elicitor is perfectly safe (Horn et al., 2012). Pain tolerance is also believed to be higher in psychopaths, a clinical population noted for low levels of fear (Miller et al., 2014).

These seem like reasonable reasons for postulating fear as the affective component of pain, but they don't withstand scrutiny. Anxiety is not the only emotion that reduces pain thresholds. Anger and sadness do as well (van Middendorp et al., 2010). There is also evidence that viewing disgusting pictures makes pain more unpleasant and intense (Meagher et al., 2001). It seems that any negative emotion can worsen the experience of pain. Moreover, when it comes to the facial recognition, pain is more frequently confused for disgust and even embarrassment than for fear (Kappesser and Williams, 2002).

Different problems arise for the Horn et al. (2002) study, which claims to show diminished negativity for safe pains. They use an indirect measure to show this: they measure negativity of pain through the startle response and they manipulate the safety of pains by making the occurrence of a painful stimulus highly predictable. But this method runs the risk of triviality. In effect, the authors show that predictable pains are not startling, which is hardly surprising. When they ask participants to explicitly report pain negativity, they find that even predictable pains are unpleasant. As for psychopaths, they have deficits with negative emotions such as sadness, in addition to fear, so diminished fear may not be responsible for their relative pain tolerance. Thus, the evidence linking fear and pain is far from decisive.

These efforts to identify an emotional correlate of pain's affective dimension have not been overly convincing. There is also a more general reason to think that such efforts will fail. The problem with this approach stems from the nature of emotions. There are many theories of what emotions are, and this is not the place to adjudicate those grand debates, but let's consider one view, which I happen to endorse (Prinz, 2004; see also Damasio, 1994). According to William James (1884), emotions are perceptions of patterned changes in the body. When an emotionally evocative situation arises, the body prepares for a behavioral response. Those preparations are the source of emotional experience. In addition, emotions carry information. In particular, they convey changes in an organism's relations to its environment that bear on well-being. For example, anger indicates that one has been mistreated or maligned, sadness

indicates loss, and fear indicates danger. These are called core relational themes (Lazarus, 1991). Thus, emotions have bodily perceptions as their form and core relational themes as their content. This picture is not universally endorsed by emotion theories, but most endorse something analogous to at least one of these two aspects of emotion.

I want to suggest that both aspects – emotional form and emotional content – can be absent in the case of pain. With respect to form, it must be granted that pains characteristically involve perceptions of the body. But that involves the sensory component of pains, not the affective component, which is our concern here. Moreover, pains don't involve global bodily patterns that relate to behavioral preparation, but rather to highly specific bodily injuries and perturbations, which vary from case to case. So pains involve the body in a way that differs from emotions. Emotions also differ from pain when it comes to content. Pains carry information about bodily injuries and perturbations, not core relational themes. The affective dimension of pain tells us something further: it presents these bodily insults as bad in some sense, but this is not the same as an organism-environment relation that bears on well-being. Pains don't indicate how an organism is related to its environment. More like hunger, thirst, or fatigue, they indicate something about the organism itself. Emotions use the body to convey information about how an organism is faring in the world, but pain (and these other states) use the body to convey information about the body itself. Thus, it seems mistaken to equate pain, or pain's affective component, with an emotion.

2.3 *Affect as negative valence*

The foregoing discussion raises doubts about the widespread view that pain's affective component is an emotion. A variety of different negative emotions can influence pain, but none of them stands out as a constituent part. Pain's negativity remains unexplained. An emotional account might be made to work in the end, but there is reason to consider other possibilities. Here I want to suggest one alternative. Rather than equating the affective component of pain with an emotion, I want to suggest that the affective component is something called "negative valence."

The term "valence" is used in the emotion literature to refer to the fact that all (or perhaps most) emotions are positive or negative. Positive emotions include joy, pride, relief, amusement, marvel, and admiration. Negative emotions include fear, disgust, anger, sadness, guilt, and shame. There is no agreed upon view about what makes an emotion positive or negative, but most researchers agree that *something* must account for such a classification. Most suggest that each emotion contains something that makes it positive or negative. That "something" is not itself an emotion. Rather, it is a component part that can occur in different emotions, and perhaps too in mental states that we would not classify as emotions. For example, hunger is not an emotion, but, like negative emotions, it seems to have negative valence. Conversely, sexual arousal may not be an emotion, but it has positive valence. Valence, then – negative valence in particular – could be

a component of pain. This would explain why pain negativity is like emotional negativity without supposing that pain negativity is itself an emotion.

To flesh out this account, we need a theory of valence. A number of theories can be found in the literature. Here I will offer a quick review, focusing on negative valence, and then I will introduce the theory I favor. For a more detailed discussion, see Prinz (2004: chap. 7). According to one account, valence is a feeling. Negative emotions feel bad. Though highly intuitive, I think the feeling account won't work. Notice that different negative feelings feel bad in different ways: the sickening revulsion that attends disgust differs from the heart pounding, respiratory unrest of fear. To say that these emotions feel bad is to say that we dislike these specific feelings; there is no phenomenal common denominator that unites them. Indeed, under some contexts, the very same feelings might be welcome: fear on a roller coaster or disgust during a graphic horror movie. Rather than equating valence with a feeling, we would do better to explain what happens when a variety of different feelings are taken as bad.

A second theory of valence focuses on behavior. In particular, negative emotions are associated with withdrawal behaviors and positive emotions associated with approach behaviors. This suggestion is more promising than the feeling theory. We can say, for example, that fear and disgust both motivate us to withdraw from their causes, even if they feel different. We can also say that fear loses its negative valence in those accessions when it is not accompanied by a withdrawal motivation: as with fear on a roller coaster for those who enjoy such experiences. Trouble for the view arises, however, when we consider other cases. Anger, for instance, is widely recognized as negatively valenced, but it is associated with an approach behavior, namely aggression (Carver and Harmon-Jones, 2009). With sadness, too, there is often a desire to approach those who might offer comfort, as when a child cries out for a caregiver. Even fear can trigger a fight response when a person is cornered. So there is no systematic relationship between negativity and withdrawal.

A third approach, which I favor, improves on the approach/withdrawal theory. Rather than identifying valence with specific kinds of behavior, we can draw on animal learning theory. In studies of conditioning, it has been found that certain stimuli (called reinforcers) increase a behavior when they are introduced, and others (called punishers) decrease behavior. To make sense of this, we must assume that reinforcers trigger psychological states that function as a kind of command: "more of this!", while punishers trigger the command: "less of this!", where "this" refers to a state that is bound to the command. In the case of emotions, "this" refers to bodily sensations, such as the respiratory constriction of fear or the giddy energy one experiences when delighted.

These internal commands lead us to seek out behaviors that would serve to obey them. It is generally difficult, however, to directly affect the duration of a sensation; we cannot easily cause a sensation to cease or continue by act of will. Therefore, the internal commands will usually be satisfied by changing behavior with respect to eliciting conditions. Thus, if we are eating delicious potato chips, the inner command that says, "more of this!" will lead us to keep eating, and if

we start to feel overly full, the internal command that says, “less of this!” will lead us to stop. These commands often relate to approach and withdrawal behavior respectively. We approach the bowl of chips when savoring their deliciousness, and we avoid it when we greatly exceed a satiation point. But the relationship to approach and withdrawal can vary as a function of context. For example, when we are hungry, the bodily state of deprivation is bound to a command that says, “less of this!”, but we satisfy that command by consuming food, which is an approach behavior. The point can also be illustrated with the examples above. Suppose you are threatened by another individual, and an internal signal says, “less of this!” That command might be satisfied by withdrawing, but it could also be satisfied by aggressing against the one who threatened you. Likewise, if you want less sadness, you might approach those who can comfort you.

It is plausible, then, that the valence of an emotion consists in one of these inner commands – what I call valence markers. A positive emotion is one that comes with a command: “more of this!”, and a negative emotion comes with the command: “less of this!” Thus we work to have and sustain positive emotions, and we work to stop or avoid negative emotions. On this view, valence is not a feeling or a specific behavioral disposition, but rather an inner marker that tells us which of our states we should try to maintain or increase and which we should try to eliminate.

Valence markers are not restricted to emotions, as the potato chip example illustrates. Gustatory pleasure is positively marked, and over-satiation is negatively marked. This account lends itself naturally to pain. Normally, we want to rid ourselves of painful sensations. We nurse our wounds and take painkillers. Pain might lead us to withdrawal behaviors (as when we leave an uncomfortable seat), or approach, as when we seek hot water to soothe aching feet. Both behaviors satisfy the “less of this!” command. We try to reduce the sensory component of pain. The overall and highly varied range of pain behaviors suggests that there are valence markers at work. Pain’s affective component, I suggest, is negative valence, and negative valence is an inner command to diminish the inner state it accompanies. In the case of pain, negative valence is a command to diminish the sensory component.

This account can be contrasted with some other theories of pain’s negativity. According to desire-based theories, pain combines a sensory component with a desire to rid oneself of that component (e.g., Armstrong, 1962). On the valence theory, pain’s negativity is explained with reference to a command rather than a desire. I think commands are a more plausible construct than desires, at least if we understand desires as they are understood in folk psychology. Desires often lack the immediacy and motivational pull associated with pain’s negativity. I might want to improve my French without doing anything about it. Valence markers are commands that encourage immediate action. In addition, we can imagine a situation where a curious person wants to have a painful sensation to know what it’s like; in that case, the person desires the sensation, but it is still negative – the sensation issues a command to stop, but that command is ignored.

The valence theory also contrasts with evaluative views (e.g., Cutter and Tye, 2011), which say that pain states include a component that represents their

sensory components as bad. To me, such views are overly semantic; they capture negativity by specifying what pain represents. The valence view also appeals to an inner representation, but that representation is a command, not an ascription of value, so it has a more direct bearing on action. It is possible that representing pain as bad can be interpreted as an injunction to act, but then it would be a notational variant on the command view.

The valence theory comes closest to so called “imperativist” theories of pain, which equate pains with commands to rid ourselves of bodily states (e.g., Klein, 2007; Martínez, 2010). One difference is that the version that I am defending is derived from an independently motivated theory of valence. Another difference, which I will discuss below, concerns the phenomenology of inner commands (or lack thereof).

Like these other theories of pain negativity, the valence theory should not be taken to imply that painful sensations are always negative. Occasionally, there are painful sensations that we don’t mind: such as the sting of exercise or spicy food. In these cases, the feeling is accompanied by something positively marked (endorphin release or gustatory pleasure), and context leads us to view the experience as beneficial, thereby changing its valence. In that case, the inner command might say, “more of this!” One attractive feature of this account is that it parallels cases of valence-switching in emotion. Those who like horror films might have an inner command that says, “more of this!”, when watching scary scenes. So-called negative emotions might be best described as emotions that are ordinarily negatively valenced. Likewise, pain is ordinarily negative, but not always.

The valence theory of the affective dimension can also explain a puzzling theoretical reversal that we encountered above. Philosophers have traditionally suggested that negative emotions contain pain as a component part, and psychologists have frequently suggested that pain contains negative emotions. This historical shift is difficult to account for, and it suggests that one group of thinkers got things completely backwards. The valence story provides both a diagnosis and a solution. Pain does not contain a negative emotion, and negative emotions do not contain pain; rather, they both contain a shared component, negative valence. The tendency to associate pain and negative emotions is explained by this overlap, and both philosophers and psychologists are partially right: negative emotions contain something that is found in pain, and conversely.

I am not aware of any explicit endorsements of the valence theory in the literature, but it fits with a long tradition of research. In animal learning theory, painful stimuli are regarded as punishers, and this makes sense only on the assumption that pain itself has an inner marker of the kind I have been describing. The valence theory also fits with the language that Melzack and Casey (1969) use when introducing the idea that pain has an affective dimension. They refer to this as a motivational dimension. As an inner command, negative valence can be regarded as a motivator: it urges us to behave in ways that satisfy the command. The term “motivational component” is therefore more apt than “emotional component,” which has become more popular.

Interestingly, Melzack himself has recognized that emotions and pains have a common component rather than supposing that one is a component of the other. In a co-authored paper, he suggests that pain contains a negative response that is also found in “fear, anxiety, stress, loss of loved objects, and other psychological states” (Loeser and Melzack, 1999: 1608). Unfortunately, he chooses to label this negative response “suffering.” Though plausible at first, this label is ultimately unhelpful. We tend to think of suffering as an emotion, and that implies, implausibly, that there is an emotion contained in other negative emotions. Furthermore, suffering is usually construed phenomenologically, which implies a phenomenological similarity between pain and these negative emotions, which doesn’t ring true. It would be better to say that each can be a form of suffering, rather than each contains or causes suffering. That use of the term leaves us with a question: in virtue of what do these phenomenologically divergent states all qualify as suffering? Valence, I submit, provides an answer. They are all negatively valenced.

In summary, I have proposed that the affective dimension of pain is not an emotion, but valence, and valence is an inner imperative that says, “less of this!” I now want to consider some explanatory fruits of this proposal.

3 Benefits of the valence theory

3.1 Pain neuroanatomy

The valence theory of pain’s affective component can account for a wide range of empirical observations. Some of these have been touched on already. Here I will elucidate some of the main explanatory payoffs.

The first payoff concerns functional neuroanatomy. We say above that the affective component of pain has been most frequently associated with activity in the anterior cingulate cortex. This pointed toward an emotion theory of the affective component, because the cingulate is a key player in emotional responses. But this is also consistent with the valence theory, since emotions have valence. Cingulate activity is a common denominator across emotions, despite their variation, making it a plausible candidate for a correlate of valence (Phan et al., 2002). We also saw that the cingulate plays a role in other functions, such as conflict monitoring and error detection, which is consistent with the valence interpretation. More tellingly, there is work suggesting that structures in the anterior cingulate contribute to reward-based decision making (Bush et al., 2002).

There have also been some neuroimaging studies directly investigating valence, as opposed to some other aspects of emotion, such as arousal. Here, again, the anterior cingulate is often implicated (e.g., Chiu et al., 2008; Schulz et al., 2009; Heinzl et al., 2010). Admittedly, valence has not been associated with cingulate activity in every study, but the inconsistency in the literature may reflect variation in methods, operationalizations of valence, and temporal resolution of imaging techniques. The overall pattern confirms that the anterior cingulate contributes to valence.

3.2 *Pain exacerbation*

The valence theory can also account for factors that increase pain sensitivity. One of the key findings was already encountered above: pain thresholds lower with the induction of negative emotions. Sadness, anger, fear, and disgust can all impact the experience of pain for the worse. This is puzzling on the view that pain contains a specific negative emotion. It could be explained on the view that pain contains a variety of different emotions, but, as we have seen, this seems implausible; it seems more likely that pain causes sadness than that pain contains sadness. The simplest explanation is that different negative emotions exacerbate pain because they each contain something that pain also contains, namely, negative valence. This may be an additive effect, wherein one has two concurrent negative imperatives originating from different sources. It could also function like priming, wherein the valence marker in the negative emotion potentiates the valence marker in the pain, making it stronger in imperative force or more accessible to information processing systems.

Negative emotions are not the only transient states that lower pain thresholds. There is also evidence that sleep deprivation has a similar effect (Lautenbacher et al., 2006). This is surprising on the hypothesis that pain has an emotional component, but it makes sense on the valence account. Negative valence is an inner command to cease and desist. This is precisely the state of mind one gets into after missing a night's sleep. The bodily state of fatigue is accompanied by a command to alter state. That command amplifies the command that arises in pain.

There is another finding that is not predicted by emotional theories of pain affect but follows naturally from the view of the valence theory: pain sensitivity increases with hunger. Pollatos et al. (2010) found that healthy volunteers had lower thresholds and tolerance for pain after two days of food deprivation. Hunger is a negatively valenced state, but it does not necessarily result in particular negative emotions. If the negativity of pain consists in negative valence, this relationship between hunger and pain is unsurprising.

3.3 *Pain reduction*

As we have just seen, the valence theory finds support in research on pain exacerbation. It is also consistent with research on pain reduction. A variety of different factors reduce pain, and some of these may be explained, in part, by a postulated reduction in negative valence.

Pain can be reduced by viewing positively valenced pictures (Kenntner-Mabiala and Pauli, 2005). The diversity of the pictures used has been extraordinary: puppy dogs, pornography, babies, loving couples, sailboats, skiers, and astronauts floating in space. There is also research showing that humor decreases pain. This has been demonstrated with stand-up comedy, situational comedy (e.g., Zillmann et al., 1993), and slapstick (Weisenberg et al., 2005), among other methods. This suggests that there is no specific positive emotion driving the effect, but rather positivity as such, which may mitigate negative valence. The exact reason for this is not fully understood. There is evidence that positive and negative valence

can operate independently (Cacioppo and Bernston, 1994), so it cannot be that positive valence markers are incompatible with negative valence markers. In principle, too, there is no reason why one couldn't simultaneously have internal commands saying "more of this!" and "less of this!", especially when they are directed at different sensations or elicitors. On the other hand, numerous studies have also shown that valence can operate in a bipolar way, with one reducing the other (Green et al., 1999). This may be how valence markers operate by default. In line with this, think about the fact that valence tends to spread. If your mood is elevated by a sunny day, you will come to have a brighter outlook towards things other than the weather as well. Valence is hard to contain. The spread of positive may also result in a suppression of negativity, which would explain why positive states induced in many different ways can reduce pain negativity. This just seems to be a basic aspect of how valence ordinarily operates.

Exercise has also been widely studied for its analgesic effects (Naugle et al., 2012). This is probably not the result of emotional changes; it has been more frequently attributed to endorphin release. Endorphins, in turn, may reduce pain by modulating the systems that underlie negative valence. For instance, research with animal models suggests that opiates, like endorphins, reduce the impact of punisher-based classical conditioning (Mauk et al., 2002). Strikingly, there is also evidence that stimulation of opioid receptors can reduce fear conditioning without reducing shock reactivity (Knoll et al., 2007). This might be tantamount to fear without negativity, which parallels the cases of painfulness without negative affect associated with morphine. This would support the conjecture that opioids are negative valence antagonists.

It is no revelation that opioids are analgesic. More surprising is the analgesic effects of swearing. People tend to shout out expletives when we sustain minor injuries, and this, it turns out, can reduce pain (Stephens et al., 2009). The valence theory offers a possible explanation. Negative valence, recall, is an inner imperative, which says, "less of this!" If so, reduction of pain negativity can come with interventions that satisfy or otherwise quell that imperative. Swearing is a symbolic action, but its normal function is akin to aggression: it is a symbolic attack on something that has been construed as a threat. Verbally attacking pain may function analgesically because it is actively doing something in response to the command. The mere act of symbolically attacking pain may be akin to the act of seeking treatment. Both effectively address the inner imperative, thereby reducing its force.

This interpretation of the swearing effect leads to the prediction that aggression itself might serve to reduce pain. Aggression is an active state that signals an organism's commitment to do something about something that threatens it. Being in this active state may appease the "less of this!" command to some degree. Research on violent video games bears this out (Stephens and Allsop, 2012). As compared to a golf game, violent shooter games increase aggression and also increase pain tolerance.

The flipside of aggression is surrender. Suppose that instead of fighting a pain, one were to acquiesce. On the valence theory, this too should alleviate symptoms.

With acquiescence, the command, “less of this!”, is rebuffed with an attitude that says, “maybe this isn’t so terrible.” A variety of acceptance-based pain therapies have been developed based on that principle. It turns out that these are effective in reducing pain (Veehof et al., 2011). Similarly, it has been said that pain tolerance goes up when people are falsely informed that a pain is beneficial (Benedetti et al., 2013). Here again, there is a motivation to acquiesce, and that proves to be analgesic. These surprising findings are predicted by the valence theory.

3.4 *The effects of pain*

The valence theory may also shed light on some of pain’s psychological and behavioral effects. Some of these have already come into view. We have just noted that pain can play a role in conditioning behavior. This is just what would be expected if pain has negative valence as a component. Negative valence is the internal equivalent of a punisher.

Pain can also lead to avoidance and withdrawal from life activities. Some of this may be due to disability, but the affective dimension of pain may also promote withdrawal behavior as a way to cope with the inhibitory imperative of negative valence. It must be added, however, that negative valence does not always result in withdrawal. As argued above, there are cases where negative states lead to approach behaviors (as when we seek social comfort in sadness). In that context, I noted that pain, too, shows variable behavioral embodiment. People in pain may seek social support or medical interventions rather than withdrawing passively. They may also try to distract themselves in an open-ended variety of ways. For this reason, the notion of “pain behavior” is almost absurd: the range is too varied to be treated as a coherent class. But, if we look beyond behavior to behavior motivation, we find unity. Pain behaviors are typically those that are directed towards pain alleviation. This may sound like a platitude, but it becomes a puzzle without an adequate account of pain’s affective component. The valence account says we aim to alleviate pain because pain contains a command that compels us to do so.

It might be objected that the valence account actually predicts a narrow range of behaviors, rather than behavioral variation. The account says that pain commands us to reduce its sensory component, so it might seem to predict that pain behavior will normally involve some effort to soothe affected areas of the body. If so, the fact that we do other things, such as soliciting social support, may appear difficult to explain. Two things can be said in response. First, people do in fact tend to soothe affected body parts when they are in pain, and this may be the most immediate reaction in most cases. Second, the theory does not entail that our coping be directed at the body. Rather, it predicts that we do anything that might deal with the inner imperative. That might include removing the noxious stimulus, distracting attention away from the sensation, or taking a pill that reduces the imperative’s urgency. Through learning and conditioning, we can develop wide ranging dispositions for pain management. It is sometimes easiest to eliminate a sensation by indirect means. By comparison, if one is uncomfortably

hot, it doesn't usually help to soothe one's overheated skin; it is easier to move to a cooler location.

In addition to behavioral effects, pain has psychological effects. We have seen that pain is exacerbated by negative emotions. Conversely, it can also cause such emotions, as when pain makes us irritated or depressed. An emotional theory of valence might suggest that this occurs because pain contains these emotions. A more plausible story – given the variability in pain's emotional effects – is that pain causes negative emotions as a function of the negative valence command. For example, protracted pain may be depressing because we come to realize that we are not succeeding in meeting the “less of this!” command. Protracted pain is a sign of a losing battle. Pain may cause anger when we feel that someone else is to blame for being in a negatively valenced state.

Paradoxically, pain may also lead to a reduction in some negative emotions. There is research suggesting that pain can reduce guilt (Bastian et al., 2011). There is also evidence that pain can reduce fear under certain circumstances (Hollin and Derbyshire, 2009). Such findings are puzzling on any theory, and their authors offer complex explanations. The authors of the guilt study suggest that pain may alleviate guilt because it makes people feel like they have been punished. The authors of the fear study suggest that we may have evolved to become fearless when in pain, because that facilitates escape behavior. But there are simpler and more unified explanations available on the negative valence theory of pain's affective dimension. First, negative valence may have an opponent after-effect, once a painful stimulus is removed (cf. Leknes and Tracey, 2008). In the animal learning literature, it is known that removing a painful stimulus is reinforcing (called negative re-enforcement), something we experience positively as relief. This may counteract some negative emotions. Second, there may be attributional effects at work. People who have just experienced pain may erroneously attribute the negativity of guilt and fear to the painful stimulus, thereby underestimating the negativity of those emotions. Third, attention may play a role. Pain's negativity demands attention, by issuing a “less of this!” command. This may attenuate the attention available for the comparable demands in these other negative emotions. These are compatible explanations, and the latter two carry a testable prediction not shared by the punishment and evolutionary accounts: pain's paradoxical capacity to diminish negative emotions should be greater when those emotions arise in a way that is unrelated to the pain. This has not been directly tested to my knowledge, but, when taken together, the literature on pain and negative emotions provides some support.

3.5 Psychological “pain”

The valence theory may also shed some light on so-called psychological pain. I say “so-called” because I think there are conceptual grounds for denying that psychological pain really is pain. This denial goes against the grain of contemporary research. Recent findings have been said to support the conjecture that

psychological pain is pain in a literal sense. I want to suggest a different possibility that may prove to be more satisfying.

The literal interpretation of psychological pain is said to find support in recent neuroimaging data. Researchers have found that psychological pain and physical pain have some shared neural correlates. In a pioneering study, Eisenberger et al. (2003) had participants play a virtual ball-tossing game, which was rigged so that they would be excluded from this enjoyable activity by other players. This induced feelings of social rejection, a kind of psychological pain. These feelings were correlated with brain activity using fMRI, and the results showed a relationship between psychological pain and activity in the dorsal anterior cingulate cortex. The authors conclude, "social and physical pain share a common neuroanatomical basis" (291). The authors use this overlap to explain why "rejection hurts," inviting a literal interpretation of the phrase "psychological pain." Eisenberger has become the leading proponent of this interpretation (see Eisenberger, 2010), and she has identified other studies that provide further support. For example, a study by Gündel et al. (2003) finds an association between grief and cingulate activity (participants view photos of deceased loved ones). Eisenberger interprets this activity as establishing the painfulness of loss.

The literal interpretation of psychological pain is not firmly established by such findings, however. It has been well-known since the early days of social neuroscience that the anterior cingulate is active in negative emotions. The fact that it is also active during experiences of physical pain does not show that negative emotions are literally painful. Rather, it shows that they share something in common with physical pain. This is precisely what the valence theory of pain's affective component asserts. Cingulate activity is a correlate of negative affect, I have suggested, not the sensory component of pain. This interpretation is also consistent with another Eisenberger study, which reports a correlation between individual sensitivity to physical pain and sensitivity to social rejection (Eisenberger et al., 2006). This is just what we should expect on a shared component model.

These considerations suggest that the presence of anterior cingulate activity is not sufficient for supporting a literal interpretation of psychological "pain." In fact, the Eisenberger findings provide reason against a literal interpretation. The fMRI studies that she marshals in support of her conclusion fail to find an association between psychological pain and activity in the somatosensory cortex, which is the other key contributor to physical pain. Without showing such activity, the comparison to physical pain is incomplete and potentially misleading.

This objection to the literal reading has been challenged by Kross et al. (2011). They conducted a study in which participants in an fMRI scanner viewed photographs of romantic partners who had recently broken up with them. When viewing these pictures as compared to pictures of friends, brain activation was shown to increase in somatosensory areas akin to those that are active during episodes of physical pain. The authors conclude that social rejection actually hurts physically. Unwanted break-ups, they say, cause physical pain. This would be a clear victory for the literal interpretation of "psychological pain."

The Kross et al. study is open to an alternative interpretation. It turns out that the somatosensory response in question has also been observed in fMRI studies of sexual arousal and romantic love (e.g., Arnow et al., 2002; Cacioppo, 2012). The participants in the Kross studies are looking at their recent romantic partners and, by the selection criteria in the study, they are still interested in those partners; they are victims of unwanted break-ups. Feelings of love and even sexual arousal would be unsurprising when viewing pictures of objects of romantic interest. Indeed, the Kross finding simply replicates these studies on love and arousal. The fact that somatosensory activation is present when experiencing romantic feelings is unsurprising, and the fact that such activation is absent in other studies of psychological pain suggests that they do not underwrite the pain of rejection here.

One can also make the point against Kross by pointing out that the somatosensory activations that he observes derive from a very different source than those involved in physical pain. In physical pain, the somatosensory system is activated by way of nociceptive cells in the peripheral nervous system. For all we know, that pattern of somatosensory activation that Kross records could never be generated by nociceptive inputs. If so, the analogy to physical pain may be greatly overstated. More generally, I think the literal view takes a rough analogy too far. Literalists are impressed by the use of pain vocabulary in psychological contexts. We talk of being emotionally scarred, psychologically wounded, and heartbroken. But these terms are clearly metaphorical. Emotional trauma can leave memory traces, but not scar tissue. I would reiterate my above objection here as well. Some metaphors for emotional pain are barely coherent. We talk about “getting our feelings hurt,” but presumably feelings are not the kinds of things that can be hurt.

Instead of taking such phrases literally, we should ask why physical pain provides such a common metaphor for psychological distress. We should ask: is there something deeper to this relationship, something that moves beyond mere metaphorical similarity? The valence theory provides an answer: physical and psychological pain share something in common, namely negative valence. In this sense, psychological pains are not merely analogous to physical pains; they have an overlapping component. Rather than pushing for a literal view, which can be misleading, we should simply emphasize this overlap. Eisenberger is right to say that social rejection shares something in common with pain, and this is important. Thus, some uses of pain vocabulary in the psychological domain are more than just a metaphor, but it should not be taken literally. This compelling middle position is entailed by the valence theory.

3.6 Individual differences

Let me turn, finally, to individual differences in pain sensitivity. The valence theory offers some perspective on these as well. We already encountered one example: psychopathic traits. Psychopaths are believed to have higher pain tolerance than non-violent comparison groups. They are also known to have reduced

negative emotionality more broadly. This is supported by abnormalities in the processing of emotional stimuli and abnormalities in emotion recognition (e.g., Blair et al., 2001; Dawel et al., 2011). There is also evidence for reduced processing in anterior cingulate and other limbic areas while viewing emotional pictures (Kiehl et al., 2001). This is consistent with the conjecture that psychopaths have a relatively broad, though not all-encompassing, deficit with respect to negative valence. The pain abnormalities may be a consequence of that.

Other research points to a relationship between pain processing and some of the so-called big five personality traits (the personality dimensions most widely studied and verified in psychology). Extroverts have been shown to have higher pain tolerance than introverts, though they also complain about pain more (Philips and Gatchel, 2000). Neuroticism, in contrast, has been associated with higher levels of pain unpleasantness (Harkins et al., 1989). These results may be interpreted in terms of valence, with extroverts having less baseline negative valence than comparison groups and individuals high in neuroticism having more.

Such findings may extend to other dimensions of personality. For example, one study associates low pain tolerance with individuals who score high on tests for “experience avoidance” (Feldner et al., 2006). Like those who are sleep deprived, these individuals may have more negative attitudes towards acute sensory states. Conversely, individuals who are disposed toward feeling hope show higher pain tolerance than others (Snyder et al., 2005), which may indicate a coping strategy for dealing with negative valence.

Future work should directly explore whether valence, as defined here, plays a role in variation among pain states and traits. There is every reason to expect confirmation of the view. Such variation is harder to explain on models that equate pain’s affective dimension with a specific negative emotion or even a family of emotions. Variation in pain sensitivity is associated with a variety of different factors, including different emotions, personality, and outlook on life. The valence theory may offer a unifying account of this diversity.

4 Objection: Aversive phenomenology

4.1 *Why does pain seem bad?*

We have just seen that the valence theory of pain’s affective dimension has a number of explanatory benefits. It offers a natural explanation of diverse variables that can influence or be influenced by pain. The theory may, however, also come with a cost. It turns out that the valence theory forces us to abandon a folk platitude about pain. That may strike some as a serious objection. Here I want to suggest that the objection can be met, and even recast as an advantage.

The objection stems from the account of valence that I have been defending, according to which negative valence is an inner command. This contrasts with a phenomenological account, according to which negative valence is some kind of bad feeling. In denying that valence is a phenomenal quality, I am suggesting that the affective component of pain is not a feeling. Now, recall that the badness of pain is supposed to be captured by the affective component. It follows from the

valence theory that the badness of pain is not a feeling. Put more provocatively: pains do not feel bad.

This is a heretical view. It flies in the face of the platitude that pains feel unpleasant, and it abandons the folk-theory according to which pain avoidance has phenomenal motives. On the valence theory, we avoid pains because they command us to do so, and that command has no attendant feeling. This may sound very odd. Isn't it obvious that pain is unpleasant? This question puts pressure on the valence theory. Perhaps I was too quick to deny that valence lacks phenomenology. Perhaps I could even say that the inner command, "less of this!", has a characteristic feeling (compare Klein, 2007). Why deny the obvious?

One could try to wiggle out of objecting by arguing that there is use of "unpleasant" that is not phenomenological. That would allow one to say, "Sure, pain is unpleasant, but the unpleasantness does feel like anything." To my ear, this sounds odd, but, in any case, it wouldn't silence the objection at hand. There is a deep conviction, I suspect, that pain has aversive phenomenology. For the sake of argument, I will assume that "unpleasantness" refers to a postulated phenomenal quality, and I want to argue that, on that interpretation, pain is not unpleasant. The point could be put in a less provocative way by saying pain's unpleasantness doesn't feel like anything, but this restatement would do little to appease those who insist that pain contains a bad feeling in addition to the felt qualities of its sensory components. This conviction cannot be dismissed, but it can be sensibly denied.

There are two compatible strategies for rejecting the thesis that pains are phenomenologically unpleasant. First, one can expose an ambiguity in the platitude that pains feel bad. This might be taken to mean that pains are feelings that we don't like to have, or it might be taken to imply that there is some feeling of badness contained in every pain. The first reading is consistent with the valence theory. Indeed, it is an entailment. If valence commands, "less of this!", then pains contain a dimension that calls for their elimination. Thus, pains are things we are compelled to eliminate, and that is tantamount to disliking them. The second reading is incompatible with the valence theory, but it is also both unnecessary and extravagant. We don't need to posit a bad feeling to explain the badness of pain, given the foregoing account of pain's negativity.

Positing bad feelings is a kind of reification. It is an unwarranted move from "pain feels bad" to "pain has a bad feeling." Such a bad feeling is an extra posit, which may actually idle under scrutiny. Suppose we ask, why do we avoid pain? Folk wisdom says, on account of its unpleasantness. But, if we construe unpleasantness as a bad feeling, this explanation only pushes back the initial question: we can now ask, why do we avoid unpleasant feelings? The answer seems obvious only if we assume unpleasant feelings intrinsically motivate their own elimination. But why assume that? Unpleasantness is usually characterized as a phenomenal state, and, conceptually, phenomenal states are independent of functional roles. It is conceptually coherent that a feeling of any kind could occur without any motivation to eliminate it. If, on the other hand, we define the badness of pain as a command to rid ourselves of pain, then avoidance motivation follows

directly from the badness. Put succinctly: valence is both necessary and sufficient to explain pain avoidance, whereas bad feelings are neither.

Against this, one might object that there is certainly something phenomenological that occurs when we are in pain, and the phenomenology in question goes beyond the sensory dimensions of pain. When pain shifts from being tolerable to being intolerable, there can be a shift in feeling, and that shift needs to be accounted for.

This objection hinges on a phenomenological intuition, which might be challenged, but rather than getting into a battle of intuitions, I want to suggest that we can accommodate the suggestion that there are phenomenal differences between tolerable and intolerable pains. We can do this by suggesting that the former have distinctive phenomenal effects (in the interest of terminological consistency). Pains often cause various feelings in us, including negative emotions. These causal effects may account for the intuition that there is a phenomenal aspect to pain's negativity. In particular, one might propose that while pain doesn't feel unpleasant, it causes a variety of negative feelings. For example, pain may cause us to feel anxious or depressed.

These strategies for explaining away the aversive phenomenology of pain gain further support when we consider the ways in which clinicians measure pain's affective component. Earlier I mentioned Melzack's (1983) McGill Pain Questionnaire, which is a standard instrument for measuring pain. The questionnaire contains a subscale designed to measure the affective component of pain, and that subscale is divided into five dimensions. One might expect these dimensions to support the folk theory that pain contains unpleasant feelings, but in fact it does not. No item listed in the subscale demands such a reading. The majority of items seem to refer to effects of pain rather than component phenomenal states.

The first dimension of the affective subscale is called "tension." It includes words such as "tiring," which is most naturally interpreted as an effect of pain rather than a phenomenal component. The next dimension, labeled "autonomic," includes items such as "sickening," which again sounds like an effect. The third dimension is called "fear," which comprises several synonyms for fear, such as "terrifying." If one were in the grip of an emotion theory of pain's affective dimension, one might interpret this term as a component feeling of pain. But it is just as natural to read this term as an effect: pain can cause terror. That may even be a more natural reading: we become terrified *because* of pain, not in and of the pain itself. The terror can come and go even as the pain remains constant.

The next dimension of the affective subscale is called "punishment." It includes terms such as "cruel," "vicious," and "killing." These might first appear to be phenomenal descriptors, but a moment's reflection reveals that they are metaphors. Indeed, they are metaphors that go hand in hand with the valence theory. Negative valence is the inner analogue of a punisher. When negative valence is high, it can be compared to a punisher that is vicious, cruel, or killing. Against this suggestion, it might be objected that the punishment associated with negative valence is not consciously felt in my theory, whereas terms such

as “cruel” are phenomenal in nature. But this objection begs the question by presupposing that there is a phenomenology of “cruelty.” To me, it is totally obscure what that phenomenology consists of beyond an intense and persistent sensory experience that I am motivated to fight. Notice that “cruel,” “vicious,” and “killing” are functional concepts. One could define “cruelty” as causing of gratuitous pain, for example. Those who describe their pains as cruel may be expressing the thought that their pains are undeserved and unrelenting. Likewise for “vicious.” When pains are described as “killing” (as in, “my back is killing me”), we again seem to have functional properties in mind: such pains debilitate and demand attention. These functional characterizations can be related to negative affect on the valence account: unrelenting pains frustrate our efforts to satisfy the “less of this!” command, and debilitating pains make that command more urgent. We can recognize these increased demands in our motivational states without having any phenomenological elements added to the intensely felt sensory elements. The metaphors on the punishment subscale in no way entail that the affective dimension is phenomenological.

The final dimension on the affective subscale is called “miscellaneous.” It includes two terms: “blinding” and “wretched.” The first refers, once more, to an effect of pain. Intense pain may indeed be blinding, since it can consume attention in a way that diminishes visual acuity. The term “wretched” is the only one on the entire subscale that sounds phenomenal and cannot be readily interpreted as an effect of pain. Still, wretchedness can be construed as an evaluative category, rather than a specific phenomenal quality. It’s something we might apply to food, say, or to a painting. When we do so, we don’t imply that there is a specific taste or visual quality that is wretchedness. Rather, we imply that we dislike the tastes and sights we are experiencing. Likewise, to call a pain wretched is to say we dislike its sensory properties. Disliking can be explained, as suggested above, in terms of valence.

In summary, the alleged unpleasantness of pain can be explained away, and the scale that is standardly used to measure such unpleasantness does not refer explicitly to the phenomenal qualities of pain rather than to its effects. The valence view is consistent with pain measurement and more parsimonious than the phenomenal theory.

4.2 Against unpleasantness

Parsimony is not the only thing that recommends the valence theory over the phenomenal theory of pain’s affective component. Let me end with two further reasons for preferring it.

Phenomenological theories of the affective component are vulnerable to an objection that was raised earlier against phenomenological theories of valence. Disgust, sadness, anger, and fear are all negative emotions, we observed, but they don’t seem to contain a shared feeling of negativity. The badness of disgust differs markedly from the badness of grief. Negative emotions lack a phenomenal common denominator. I would say the same of pains. We have many different

kinds of pains: sharp, burning, dull, and so on. It is far from obvious that they feel alike in some way. A razor burn and a bellyache seem to be phenomenally disjoint. Their phenomenology seems to be fully captured by their divergent sensory components. Notice, correlatively, that when we experience pains, we can easily focus attention on minute details of their sensory qualities: their spatial distribution, their pulsation, pressure, and itch. It's almost as if we could visualize these sensory qualities: some pains are faint nebulae, and others are like spikes or flashing bolts of light. But we can't bring the affective dimension into such attentional focus. Once these other features are described, there is no intrinsic quale left that we can't describe. This supports the valence theory.

Another argument against aversive phenomenology can be extrapolated from work on analgesics and other interventions that block the affective component of pain without blocking the sensory component. When asked to report on their experiences, people in these dissociative states have no difficulty recognizing pain as such. For example, a patient who was given a lobotomy for chronic pain testified that her pain is "still there. I just don't worry about it anymore ... In fact, it's still agonizing. But I don't mind" (Yancey and Brand, 1997: 2010). Another patient in a study of the analgesic effects of nitrous oxide explained that, "Pain was as bad with the nitrous oxide, but did not bother so much" (Chapman et al., 1943: 874). Similarly, Franklin (1989) reports that patients give equally high pain scores when on morphine, and they say that their pains "hurt," but they also say they are not bothered any more, and they do not increase their morphine intake when given the opportunity. These first-person characterizations suggest that people who lack an affective component have no difficulty recognizing their pains. They continue to describe them as painful, as hurting, and even as agonizing. If the painful phenomenology of pain were located in the affective dimension, we would expect a recognition impairment in these cases. We would expect the affectively anaesthetized individual to report pains as if they were merely touch sensations, such as tugging or pulling or pressure. By analogy, imagine a drug that removed the hedonic effects of tickling. People on this drug would not say, "that tickles, but it doesn't make me laugh." They would more likely, say, "I can feel you touching me, but it doesn't tickle."

The remarks of these patients suggest that the affective dimension of pain can be removed without depriving pain of its characteristic phenomenology. I don't mean to imply that there is no change in their overall phenomenology. They are clearly less anxious about their pains; they calmly smile (Yancey and Brand, *ibid.*). The simplest explanation is that the intrinsic qualities of pain remain more or less the same, but the negative effects of pain diminish. This is what they mean when they say their pains don't bother them. The valence theory predicts that people who lose pain affectivity will continue to recognize pain. Phenomenal theories of the affective component seem to predict a more dramatic change in pain reports and pain recognition.

These are not decisive objections. They rest heavily on phenomenal intuitions and reports – including the reports of people on powerful drugs. But they weaken the case for aversive phenomenology. Aversive phenomenology is often

taken to be an obvious and non-negotiable component of pain. Here I've tried to suggest that such faith is overblown. The fact that pains strike us as bad can be explained without aversive phenomenology, and it's unclear what that phenomenology is supposed to consist in. The analgesia cases suggest that pain is readily recognized without negative affect; this is most easily explained by the conjecture that negative affect makes no direct contribution to pain phenomenology. Other interpretations of these findings are possible, of course. I am content to conclude that the existence of aversive phenomenology is far less firmly established than it might initially appear. Indeed, phenomenal theories of pain's affective component may raise more puzzles than they solve. Therefore, we should be open to the possibility that pains are not unpleasant. Crucially, this conclusion doesn't threaten the thesis that we dislike pains. Absent anesthesia, pains compel us to pursue palliation. We also find pains depressing, frustrating, and frightening, and we work hard to eliminate them. These effects and efforts are often felt even if pain's negativity is not.

5 Conclusion

For almost half a century, pain researchers have recognized that pain has an affective dimension, along with a sensory dimension. The foregoing has been an inquiry into the nature of that affective dimension. The prevailing view, though often implicit, seems to be that pains contain negative emotions. This, I suggested, may not be right. Instead, pains share something in common with negative emotions: they are negatively valenced. On the view defended here, negative valence is a kind of inner imperative. It says, "less of this!" I defended the valence theory by suggesting that it has certain advantages over emotion-based theories, and it bears many explanatory fruit. The arguments are not demonstrative. Indeed, if negative emotions contain negative valence, then the two approaches make many of the same predictions. But pain interacts with many variables, including a range of emotions as well as states and traits that do not qualify as emotions. Valence offers a tidy explanation. The valence account has an odd implication, however: it abandons the folk platitude that pain is unpleasant. This is not a reason to abandon the view. Instead, it may lead to new insights about the badness of pain.

Acknowledgments

I am indebted to Jessie McCormack for supremely helpful suggestions and corrections on an earlier draft, to Michael Brady and Jennifer Corns for invaluable feedback, patience, and encouragement, and to David Bain, whose detailed comments led to improvements on every page.

Part III

Deviant pain

7 The unpleasantness of pain for humans and other animals

Adam Shriver

Introduction

For roughly the past 50 years, philosophers have been puzzled by reports of people who claimed that they were “feeling pain” but not “bothered by it.” These reports are surprising because on one common interpretation, pain is essentially bothersome; that is, one cannot have a pain without having an experience that is unpleasant or aversive in some way. Reports of pain without unpleasantness are important for ethical purposes, since pains lacking unpleasantness are less morally relevant on many views, and also for the philosophy of mind, since many theorists have wondered if there’s an essentially motivational component that constitutes at least part of the experience of pain.

One well-supported explanation of these reports is that pains have at least two distinct components: a sensory/discriminative component and an affective/motivational component¹ (Melzack and Casey 1968). As will be discussed below, many experiments have shown that these two dimensions can be modulated independently of one another. As such, we can view the reports as instances where the affective dimension of pain is selectively impaired but the sensory dimension is operating normally.

Research on nonhuman animals has also purported, in many cases, to show a similar dissociation between the affective/motivational and sensory dimensions of pain. Due to the fact that many invasive experiments performed on other animals are not performed on humans, the possibility that these experiments are capturing the same dissociation would have important implications for understanding pain in humans, since this connection would make available a host of information, particularly at the level of individual neurons and neural circuits, about the dissociation that would otherwise be lacking.

In this essay, after reviewing the importance of understanding the affective dimension of pain for various subdisciplines of philosophy, I will describe what is known about this dimension of pain in other species and suggest how this knowledge could impact our understanding of pain in humans. Ultimately, I will argue, the main impediment to progress in understanding the unpleasantness of pain is the lack of a sufficiently detailed investigation of cases where humans report feeling pains without finding them unpleasant. In reviewing the literature on pain in animals, I do not intend to suggest that current invasive and/or painful

research is justified; indeed, it may be the case that what we learn from such research is enough in itself to provide an indictment of certain research practices. However, I do think we should use the information we currently have to improve our understanding of how pain relates to important philosophical issues.

1 The sensory/affective dissociation and why it matters for philosophers

Philosophers studying the nature of the mind (Ryle 1949) and the relevance of pain for ethics (Hare 1964) have been aware of the dissociation between pain sensations and unpleasantness since at least the middle of the 20th century. Though important empirical advances have been made in understanding pain since that time, many of the same conceptual questions remain. In this section, I describe the relevance of the affective/sensory dissociation in pain for value theory, applied ethics, and the philosophy of mind.

Before highlighting the intersection with philosophical issues, however, I have to say a little more about our current understanding of the different dimensions of painful experience. The sensory/discriminative dimension of pain, linked to activity in the primary and secondary somatosensory cortex, is generally described as consisting of three elements: the “quality, intensity, and spatio-temporal characteristics” of the pain sensation (Rainville et al. 1999). When we experience a typical acute pain, we experience it as occurring in a particular part of our body and having certain temporal properties (for example, a “throbbing” pain rapidly fluctuates in intensity). It is also typically represented as a particular kind of pain: a “cutting pain,” or a “burning pain,” or a “pinching pain,” for example. These different types of pains typically result from the activation of nerve fibers in the skin, called nociceptors, that respond to different forms of potentially damaging events, such as extreme temperatures, chemical substances, or tissue damage. The representation of these types of pain is the “quality” of the pain referenced in the quote above. Finally, pains vary in degrees of intensity. Some pains, of course, are stronger than others.

Somewhat surprisingly, however, the perceived intensity of a pain does not fully determine the immediate experienced unpleasantness of the pain. Under the right conditions, more intense pains can be less unpleasant, and less intense pains can be more unpleasant. If a person is expecting a pain, this can influence the perceived unpleasantness of a pain, even as the ratings of pain intensity stay the same (Sawamoto et al. 2000). Likewise, increased anxiety causes pains to feel more unpleasant even when the ratings of intensity stay the same (Ploghaus et al. 2001). Moreover, certain drugs appear to selectively influence the stated unpleasantness (Keats and Beecher 1950). And lesions to particular brain areas have resulted in patients who claim that they feel pains but no longer find the pains unpleasant. Frontal lobotomies (Freeman and Watts 1946, Koskoff 1949), cingulotomies (Foltz and White 1962, Corkin and Hebben 1981), and insula/parietal operculum lesions resulting from strokes (Grahek 2001) have all been associated with such reports. Most recently, research from Pierre Rainville’s

lab (Rainville et al. 1997, Rainville et al. 1999) has shown that the affective dimensions of pain can be manipulated independently of the sensory dimension, and that two distinct activation patterns during brain imaging correspond to the different dimensions.

What the above evidence has shown is that people's reports of pain intensity can diverge from their reports of unpleasantness. So if we accept patients' ratings of pain's unpleasantness, we should acknowledge that the two dimensions can come apart, at least to some degree. Nevertheless, there still are many questions remaining about the nature of unpleasantness. What role does unpleasantness, as opposed to other aspects of pain experience, play in our behavior? How was it useful from an evolutionary standpoint? Is there some other psychological capacity or behavioral indicator that could help us to understand what unpleasantness is actually doing?

Many pain researchers use the term "affective-motivational" to describe the pathway involved in unpleasantness, which implies that this pathway also underlies the motivational urge associated with pain. Some have suggested the unpleasantness of pain can be linked with the motivational force of pain²; that is, if we find a pain unpleasant, it is at least plausible that we necessarily have some motivational urge to escape the pain. However, as will be discussed below, the science relevant to this claim is currently inconclusive, and thus it remains an open possibility that the affective and motivational aspects of pain are two separate components.

Beyond motivation, there are a host of other features which may or may not be correlated with the affective dimension of pain. Austen Clark has listed all of the following conceptually distinct variables as possible features linked to the affective dimension of pain: the desire to end the pain, the feeling of urgency to do something about it, the degree to which pain grabs one's attention, the extent to which pain alters one's preferences, the degree to which the reduction of pain can serve as a reward, and the degree to which the pain reduces the probability of future behaviors (2006, p. 185). All of these are potential correlates of pain's unpleasantness that could potentially be used in third-person assessments of suffering and therefore, as will be discussed below, could be important for answering questions in applied ethics.

Understanding pain's sensory/affective dissociation is also important for various issues in the philosophy of mind. In brief, any theory of phenomenal consciousness will need to account for the full range of phenomenal experiences. Many accounts that seem to have been designed to capture aspects of conscious perception have difficulty capturing the unpleasantness of pain (and possibly of other bodily sensations such as itches, tickles, etc.). Perhaps the most prominent example is representationalist accounts, which suggest that phenomenal character supervenes on intentional content. On standard formulations of these views, any difference in phenomenal character between two states must result from a difference in the way the world is represented to be. Pains, as described above, would seem to challenge this supervenience claim, since the phenomenal character of experience might vary as unpleasantness varies without anything about

the world being represented differently (recall that the sensory features can be manipulated independently of the affective features).

A number of proposals have been put forward to characterize the phenomenology associated with the affective dimension of pain (Tye 2006, Cutter and Tye 2011, Aydede and Guzeldere 2005, Aydede and Fulkerson 2014), including the proposal to treat the intentional content of pains as consisting of imperative rather than declarative content (Klein 2007, Hall 2008, Martinez 2011, but see Bain 2011 for a critique).³ On these latter views, instead of representing the world as being a certain way, e.g., “There is a red apple on the table,” pains (or at least the affective element of pain experiences) might be better characterized as giving a command such as: “Move your hand away from the flame!”

It is not my intention in this essay to adjudicate between these different theories, but I do hope to make clear that empirical research will be important for sorting many of the theories out. After all, though there are limitations on what empirical research can tell philosophers, we can at least expect it to show which psychological features are reliably correlated with the unpleasantness of pain, including potentially interesting features such as motivational signals to avoid the pain, learning signals that change future behavior, and shifts in attentional states. Knowing these correlations will be helpful in sorting out the competing theories.

Another area where the affective/motivational component of pain is extremely important is ethical theory, both at the level of applied ethics but also regarding questions of what ultimately has value (or disvalue). Regarding the latter, many philosophers have made claims that relate to situations of “felt pain without unpleasantness” and have interpreted the cases in radically different manners. For example, in his book *On What Matters*, Derek Parfit (2011) writes: “When we are in pain, what is bad is not our sensation but our conscious state of having a sensation that we dislike. If we didn’t dislike this sensation, our conscious state would not be bad” (p. 54).

Parfit contrasts this view with that of Korsgaard, who suggests that desiring something for its own sake can confer value on that thing. Specifically about pains, Korsgaard (1992) writes, “I am not denying that when we are in pain part of what is going on is that we are having sensations of a certain character. I am however denying that the painfulness of pain consists entirely in the character of those sensations. The painfulness of pain consists in the fact that these are sensations that we are inclined to fight” (p. 103).

Though neither Parfit nor Korsgaard explicitly cite dissociation studies in their texts,⁴ these different views lead to contrasting interpretations of the pain dissociations. For Parfit, when one “has a pain but does not find it unpleasant,” the person is having a sensation of pain but does not dislike the sensation. Having an experience one dislikes is enough to make the event one of disvalue, independently of whether one also has a desire to be rid of the experience. In contrast, on Korsgaard’s view, the desire is essential for pain’s disvalue, and as Guy Kahane (2008, p. 328) points out, the interpretation of lobotomy patients as feeling the sensation of pain without being motivated to do anything about it is often cited as a reason to endorse the view that things have value because we desire them.⁵

Needless to say, since Parfit and Korsgaard are debating what confers value in the world, sorting out the relevant facts about pain matters very much (see Brady, this volume, for a related discussion). Many philosophers will note that intense suffering is bad when discussing the relationship between conscious experiences and value. Yet neither our intuitions nor the relevant science are clear regarding whether the disvalue of such experience is tied to a “disliking” reaction, to a desire reliably caused by the experience, or to some other feature of pain.

We cannot fully sort out the nature and relevance of pain’s unpleasantness through science, and philosophers have an important role to play, but as noted above, we can at least learn from the sciences which features are correlated with pain’s unpleasantness. And I suggest we will need to determine these correlations before we can feel confident in our resolution of these competing claims about the essence of the disvalue of pain.

Turning to applied ethics, understanding the unpleasantness of pain is also crucial for answering various difficult questions about the moral status of beings incapable of self-report on their conditions. The most obvious example, perhaps, is nonhuman animals. Even single-celled organisms are capable of avoiding potentially harmful aspects of their environment, but we can further subdivide beings capable of avoiding danger into those that can suffer as a result of their experiences and those that cannot. However, our ability to successfully do this will depend, at least in part, on our understanding of the neural mechanisms underlying suffering.

For example, fish scientist James Rose (2002) has suggested that the fact that fish lack the brain areas involved in mammalian pain affect excludes them from moral consideration. Rose put his arguments in terms of consciousness in general; however, the areas he cited as crucial were specifically those that are associated with unpleasantness in humans. Craig (2009) has suggested that the insula (a structure involved in the affective dimension of pain) integrates bodily information (including information about pain) into a conscious representation of the body, and argues that only humans and possibly higher primates have the relevant brain structures to support conscious awareness. And Allen et al. (2006), Shriver (2006), and Farah (2008) have noted that all mammals appear to have many of the relevant brain areas underlying the affective dimension of pain in humans. Clearly, a better understanding of pain’s unpleasantness and the underlying neural activity will have extremely important implications for our moral obligations to other animals. The relevant science will be discussed more below.

Similar considerations apply when we consider the development of human beings. Opponents of abortion often claim that fetuses can feel pain at 20 weeks. However, these claims are typically based on behaviors that are generally considered to be merely reflexive, and expert panels of scientists have suggested that it is unlikely that pain perception is possible any earlier than 29 weeks, when the relevant connections to the brain are initially formed (Lee et al. 2005).⁶ But what ultimately matters, at least for this aspect of the debate, is when in human development the *unpleasantness* of pain is first present. Pain sensations without the unpleasantness do not seem to add any new moral demands.

One final example of the significance of the affective dimension of pain for those unable to self-report: in recent years, neuroscientists have found that some patients previously believed to be in vegetative states were able to follow commands while in an fMRI scanner in a manner that could only be detected by observing their brain activity⁷ (Owen et al. 2006). A subset of these patients could also answer yes or no questions while in the scanner (Monti et al. 2010). In one such case, the patient was asked if he was in any pain; however, most of the patients able to follow commands are not able to answer yes or no questions. For these patients, it would be very useful to be able to assess if they were suffering from chronic pain in their current condition. Searching for precise biomarkers associated with the affective dimension of pain, rather than pain per se, would be an important advance in our ability to responsibly take care of such patients.

The affective dimension of pain is also relevant for many issues that don't involve patients incapable of communicating. To cite one example, bioethicists have recently been discussing the significance of a neural signature of pain, that is, a biomarker that reliably indicates the presence of pain independently of self-report. It has been suggested that this neural signature could help resolve legal disputes, such as whether a particular workplace accident has left an employee with chronic pain, and could help to diagnose whether people requesting pain medication genuinely need it or are simply attempting to get high. Again, what is relevant for such disputes, insofar as we are concerned with people's quality of life, is a neural signature that detects the affective aspects of painful experience, rather than the sensory aspects (Shriver 2014).

So, as I hope to have demonstrated, getting a clear picture of what we know about the unpleasantness of pain is extremely important for a number of issues in philosophy. I now turn to research on nonhuman animals that could help to reveal important aspects of this dimension.

2 Searching for the affective dimension of pain in other species

Since pain research on nonhuman animals has been a large target of research, one might hope that there would be a great deal of research to draw upon to learn about the unpleasantness of pain. Unfortunately, however, due to an emphasis on easily reproducible measurements, much of the research has focused on measurements that tell us very little about the affective dimension of pain. Several researchers (Vierck et al. 2008, Roughan et al. 2014, Gregory et al. 2013) have highlighted the fact that many preclinical studies involving animals rely on reflexive responses to pain controlled by neurons in the spinal cord rather than the brain, and as such have done a very poor job predicting the efficacy of various treatments on humans' experiences of pain, and particularly the affective dimension of pain.

Relatedly, many ethicists have referenced the fact that many animals, when encountering noxious stimulation, display a similar behavioral repertoire to humans', and infer from this fact that the animals are in pain. Though in broad strokes I think that the observation of behavioral similarities combined with a

principle of erring on the side of caution gives us strong reasons to treat animals as moral subjects, many of the behaviors cited would be vulnerable to skeptical critiques without the support of a precautionary principle. Common examples of behavioral indicators of pain include such actions as crying, moaning, wincing, and withdrawing from the source of the stimulation. But as it turns out, quite a few of the behaviors displayed in any particular instance of pain are not necessarily connected with the painful experience. For example, humans in vegetative states will sometimes cry out, withdraw, and/or exhibit facial contortions after noxious stimulation (Laureys 2007), but these reactions are known to be reflexive. Similarly, many of the behavioral measures of pain in tests on animals, such as tail flicks and paw withdrawal, are known to be mediated by neurons in the spinal cord.

So many of the behaviors most directly associated with pains do not seem to require the experience of unpleasantness (or pain, for that matter). At this point, the neuroscience of pain in humans becomes relevant. If we can understand the mechanisms that explain pain behavior in the absence of affect, and if we know a lot about the processes in the brain that are involved in conscious pain processing, we can shift our attention from behaviors that do not require unpleasantness and identify other activity that can be examined in more detail to try to determine the essential mechanisms of affective processing.

In humans, the two brain areas that are most consistently activated during a wide variety of pains in imaging scans are the anterior cingulate cortex and the insula cortex (along with the adjacent parietal operculum, which appears to be functionally continuous with parts of the posterior insula). Single-unit recordings in humans performed during surgeries, which measure the activity of individual neurons, demonstrated that the anterior cingulate contains neurons that selectively respond to painful stimuli but not other forms of somatosensory stimulation (Hutchison et al. 1999). EEG recordings have also found that the posterior insula responds to painful stimulation (Garcia-Larrea 2012), and direct stimulation of regions of the insula results in pain reactions in humans (Ostrowsky et al. 2002). Research has also shown that selectively enhancing or diminishing the unpleasantness of pain also enhances or diminishes activity in these areas (Rainville et al. 1997, 1999). Moreover, patients with lesions in the ACC or insula have reported feeling pains but no longer being bothered by them (Foltz and White 1962, Corkin and Hebbin 1981, Schilder and Stengel 1928, Berthier et al. 1988). Thus, there is good *prima facie* evidence that both of these regions play a role in the unpleasantness of pain, or at least of some pains.

It should be noted that for the ACC and for the insula, there are complications with any claims that the area is a “locus of pain’s unpleasantness.” For one thing, both areas (at least broadly speaking) are involved in a huge number of other processes and emotions, and there are a number of competing hypotheses about what their precise role in cognition is. This worry can be diminished somewhat as the localization claims are made a bit more precise; for example, “sadness” is often suggested to be localized in the ACC but actually tends to occur in an area in front of and below (anterior and inferior in neuroscience speak) the area most

involved in pain (Vogt 2005), and similarly the pain processing area of the insula can be kept distinct from other areas of the insula, including those that process gustatory information (Nieuwenhuys 2012). Nevertheless, at the level of grain detectable by our current best brain imaging techniques, there are no chunks of neural tissue exclusively activated by pain or by unpleasantness.

The fact that no large-scale brain regions are devoted exclusively to pain need not diminish the importance of the ACC and the insula, however. Even if there are not areas large enough to be detected by brain imaging devoted exclusively to pain, there may nevertheless be individual neurons and neural circuits that are, as was suggested from both recording and stimulating individual neurons in humans. If there are neurons or circuits within a more general brain area devoted to pain processing, then lesions to those entire areas will also destroy those neurons, and as such the behavioral effects associated with those lesions can still provide us with information about certain features of pain.

Since there is good evidence that parts of the insula and ACC are playing some important role related to the unpleasantness of pain in humans, how could it be assessed whether these brain areas, or analogous areas, are playing a similar role in other animals? Given the Problem of Other Minds and the specific difficulties of assessing mental states in nonlinguistic animals, there is of course no perfect answer to this question. However, by combining behavioral measures with knowledge of neural similarities and responsiveness to certain drugs (such as known analgesics for humans), we can generate at least plausible answers.

A converging set of evidence has suggested that the anterior cingulate cortex is involved in at least some broadly speaking affective aspects of pain perception in other mammals. As noted above, studies relying on spinally mediated behaviors have tended to be very poor predictors of the clinical efficacy of pain drugs. However, several authors have presented data arguing that a more successful method can be used based on measurements of escape and avoidance.

Sufka (1994) developed one of the earliest avoidance-based operant conditioning models of the effectiveness of analgesics. On his model, animals exposed to repeated noxious stimuli are confined to one location while under the influence of a particular drug, and then are confined to a different area, still exposed to the same noxious stimulation, while not under the influence of the drug. The animals are then offered a series of choices between the two locations, and the outcomes of the choices are measured. The rationale is that if an animal experiences more pain when not under the influence of the analgesic, it will form a negative association with the location, and as such will be more likely to choose the location paired with drugs that are effective analgesics than the other location. Of course, one worry is that because many analgesics also have positive reward properties independent of any pain relief, these tests might simply be measuring the rewarding effects of the drugs, rather than pain relief. However, Sufka later (Roughan et al. 2013) cites a number of studies demonstrating that the conditioned place aversion effects of pain relief can be dissociated from the place preference caused by the reward value of the drugs, suggesting that a separate mechanism is involved in learning to avoid aversive stimuli.

The efficacy of conditioned place preference tasks, however, depends on a crucial assumption, namely that there is a link between the unpleasantness of pain and learning to avoid the pain in the future. Unquestionably, these two concepts are distinct from one another, whether or not they are reliably correlated. As such, it is difficult to evaluate how successful this approach is at indicating the existence of the unpleasantness of pain in other animals.

For this reason, several authors have expanded upon the conditioned place preference task by also adding an escape component that more closely captures the immediate feeling of unpleasantness (as opposed to a future motivation to avoid cues that were previously paired with noxious stimuli). On these models (Vierk et al. 2008, Morgan et al. 2008, Gregory et al. 2013), animals are offered an opportunity to escape from a potentially noxious environment by moving to a new location. In contrast to reflexive behaviors such as crying out, tail flicks, immediate paw withdrawal, etc., the assumption in these studies is that the animal has to recognize the negative stimulus and then use this information to make a decision to move to a new location. Again, of course, this behavior does not prove the existence of unpleasantness or even conscious pain, since it's possible that the animals have unconscious mechanisms for motivating behavior (even complicated behavior) away from threatening features of the environment.

Interestingly, however, the effects of opiates and targeted lesions in the ACC on escape and conditioned place preference in non-human mammals mirror the dissociation observed in humans. Lesioning the anterior cingulate results in mice who still withdraw from painful stimuli, but no longer choose to escape from areas where the pain should be greater nor learn to avoid these areas (LaGraize et al. 2004). Similarly, giving the rats opioids produces the same dissociation (LaGraize et al. 2006). Thus, there's at least a *prima facie* argument by analogy that in most mammals the affective pain pathways are operating very similarly to how they operate in humans, since damage and activation of similar brain regions results in similar behavioral profiles (Shriver 2006).

3 Connecting nonhuman and human studies of pain

In some ways, though, there's an important disanalogy between the results on rats and mice and those on humans. In humans, cingulotomies produce reports of "feeling pain but not minding it" in the case of pains that existed prior to the operation. However, several researchers have found that cingulotomy patients still react as vigorously to new pains (Davis et al. 1994). But in the case of the mice, the lesions to the cingulate appear to influence their behavior in response even to new pains.

Similarly, some have suggested that a condition known as "pain asymbolia," which in humans typically involves damage to the posterior insula and parietal operculum, might be the "best" dissociation since it suggests a lack of affect even for new pains. Nikola Grahek's book *Feeling Pain and Being in Pain* (2001) presents the most detailed case for the centrality of asymbolia for understanding pain affect. As described in the book, asymbolics are still able to recognize painful

stimulation, but no longer withdraw from pinpricks, pinches, or other noxious stimulation. In fact, they often seem to approach the noxious stimuli, and have been reported to spontaneously laugh during noxious stimulation. Moreover, there are at least some reports that they fail to respond to additional threatening stimuli, such as an approaching lorry (Hemphill and Stengel 1940, p. 256).

Older studies from Schilder and Stengel (1928) and more recent studies from Marcelo Berthier's lab (Berthier et al. 1988, 1990) represent the core evidence for pain asymbolia and describe a small number of subjects with insula lesions who experienced a complete lack of pain affect despite reporting the sensation of pain. Studies by Joel Greenspan's group (Veldhuijzen et al. 2010, Greenspan et al. 1999, Greenspan and Winfield 1992) also associate posterior insula and parietal operculum lesions with decreased pain affect, but in contrast to Berthier's studies do not describe patients completely lacking pain affect. Rather, the patients in these studies show increased, but not unlimited, pain tolerance.

Research on monkeys appears to produce findings similar to those in asymbolia. First, researchers have identified, via single unit recordings, neurons in the insula that mediate monkeys' willingness to escape noxious stimuli (Dong et al. 1994). This same laboratory then performed an experiment demonstrating a dissociation between the monkey's ability to discriminate painful stimulation and the monkey's motivation to escape from the stimulation, and linked the latter again to neurons in the insula/parietal operculum region (Dong et al. 1996). Researchers in these experiments trained monkeys on two tasks. In one task, the monkey, after being prompted by a cue, was to hold down a button for a set period of time while heat was being applied to the monkey's face. The monkey could "escape" at any time by releasing the button, which would result in a rapid cooling of the thermode, but if the monkey successfully held the button for long enough, it would receive a reward. On the other task, the monkey was presented with a rapid downshift of thermal stimulation on the face. The downshift was either from a nonnoxious degree of heat to a lower nonnoxious degree of heat, or from a noxious degree of heat to a nonnoxious degree, and the monkey's task was to indicate which of these scenarios had just occurred. What the researchers found was that when the parietal operculum region of the monkeys was damaged, the monkeys could still discriminate between the noxious and nonnoxious temperature shifts in the second task, but their tendency to escape in the first task was almost entirely eliminated. Thus, these studies suggest that, similar to the dissociation in humans, the insula seems to be involved in a desire to escape the stimulus that is independent from at least some aspects of the initial pain detection.

As such, I think research combining knowledge of the role of the cingulate and insula with that of escape and avoidance behavior appears to be roughly along the right track. Even so, very important questions remain. In particular, it's not clear what brain activity is really necessary for pain affect. One interpretation of the available data is that activity in the network formed by the cingulate and insula, or perhaps one of those brain areas individually, appears to mediate the unpleasantness of pain. One could then claim that only animals that have these

brain areas feel the unpleasantness of pain in a manner similar to humans. This, however, would lead to the counterintuitive result that only mammals truly have the affective component of pain, since these brain areas have not been found in other species. Even more restrictive is the view of Craig (2009), who proposes a model of insula activity whereupon the anterior insula is responsible for generating conscious experience. Craig claims that similar regions to this particular part of the insula do not exist even in species as similar to humans as monkeys.

Alternatively, one might instead suggest that there are analogous areas that operate in a similar enough manner to the cingulate and/or insula in other species to generate unpleasantness. On this interpretation, the human cingulate and insula have specialized to perform additional operations in humans, but still play a role in the experience of pain that is played by other brain regions in other species. This is plausible, since it appears the insula involves a specialization of pain perception processes that exist in a number of other species (Garcia-Larrea 2012, Nieuwenhuys 2011). Moreover, many other species, including birds, fish, reptiles, and possibly even crabs, have been shown to demonstrate conditioned place preference behaviors.

Nevertheless, we are faced with the problem that neither any one particular measure of behavior, including conditioned place preference tasks, nor the involvement of any particular brain region, such as the insula, has been shown to reliably indicate the presence or absence of the affective dimension of pain. Only by learning more about the specific way neural activity underlies behaviors associated with pain affect can we hope to abstract the processes involved in causing pain's unpleasantness. And this, in turn, will require far more research aimed at conceptually understanding the nature of pain affect in humans. For example, though there have been a number of reports of people who "feel pain but don't mind it," most notably in cases of pain asymbolia, there has ultimately been very little investigation of what the full behavioral profile of such patients is, and how reliably this profile is linked to damage of specific brain areas. If we could discover that pain affect, but not sensory pain, was necessarily linked to a particular type of learning, or to a certain form of attentional process, this could give us a better target for understanding what precisely needs to be sought out in other species. In particular, as I will discuss in the next section, if it was demonstrated clearly in humans that the unpleasantness of pain is reliably correlated to escape impulses or conditioning, this would make an important difference in how confidently we could interpret the results from other mammals.

It is only after we have detailed knowledge of the correlations that the real questions of philosophy of mind come in. For example, let's say we discover that self-reports of unpleasantness in humans turn out to be perfectly correlated with a particular type of neural activity and with a particular behavioral profile. Even then, we can question how we should view the central processes involved. Does the particular type of neuron involved in the activity matter, or does it only matter that certain information was transmitted? Can aspects of the phenomenal character be explained by describing the activity of the neurons, or are we left with only a correlation that is not fully understood? These questions, I believe,

can only be fully addressed *after* a detailed scientific investigation that narrows down the true nature of unpleasantness by precisely determining the neural and behavioral features that reliably correlate with human reports of unpleasantness.

As such, though I think philosophers can still usefully discuss these issues in the philosophy of mind and in ethics, the conclusions that can be reached are currently limited by the absence of crucial information from the sciences. My belief is that philosophers' most useful role in this context at the present, aside from producing arguments about how we ought to interpret limited and uncertain information, is to help think of experiments conceptually clarifying the nature of pain's unpleasantness, with an eye towards addressing issues of pressing philosophical and social significance. Those questioning whether philosophers can contribute something of value to this enterprise need look no further than the poor success rate of previous animal models of pain. Having a better theoretical perspective on the nature of pain's unpleasantness will be useful in designing experiments better able to capture the features that matter.

I see philosophers and scientists working together as follows: philosophers can be helpful in keeping the focus on important theoretical questions about the nature of experience and the nature of (dis)value, and on how the findings of empirical research can inform a number of practical ethics questions. However, the theorizing of philosophers is limited in at least two significant ways that require contributions from the sciences. First, we don't yet know how important psychological concepts are related to one another, and how they should best be "lumped" and "split": is the unpleasantness of pain always accompanied by a motivational signal, or by some other relevant psychological feature? And second, we don't know how the absence of various features would influence people's assessments of their own pains: if we were to selectively inhibit the motivational system of a person, how would that influence her interpretation of pain? Thus, I see progress in understanding the unpleasantness of pain as depending upon an ongoing conversation between cognitive scientists and philosophers.

To help illustrate this idea, I turn now to two examples that I think might be especially promising for our understanding of the unpleasantness of pain: the learning signal generated by painful experiences and the motivational component of pain. These examples, I hope, will demonstrate the potential of pain science to help answer some of the questions raised above for applied ethics, axiology, and the philosophy of mind.

4 Is unpleasantness connected to learning and/or motivation?

Many theorists have proposed that there is an important connection between pain and learning. Philosopher David Bain (2013), for example, has suggested that a virtue of his account and Colin Klein's (2015) account of pain asymbolia is that they can account for why pain asymbolics have trouble learning which situations might harm them. Neuroscientist Robert Gereau has written that, "the similarities in the cellular mechanisms underlying pain and learning and memory are striking" (2008, p. 87).

Colin Allen (2004) has also suggested that exploring the relationship between pain and learning might be especially important for understanding the role conscious pains have played in evolution. Allen, however, cautions that there are also many forms of learning from noxious stimulation that apparently do not depend on conscious pain perception (see Grau et al. 2002 for several examples). Allen points to important work by Clark and Squire that found differences in the requirements for trace and delay conditioning. In a task involving delay conditioning, where a puff of air is blown on subjects' eyes after the onset of a tone but before its termination, human subjects were conditioned to blink after the onset of the tone regardless of whether or not they were aware of a relationship between the two stimuli (Clark and Squire 1998). However, in a task involving "trace conditioning," where the air puff was delivered only after a time delay following the end of the tone, only those people who were explicitly aware of the relationship were conditioned to blink after hearing the tone. Allen suggested that the apparent need for explicit awareness in humans for trace conditioning might be a useful tool for searching for consciousness in other species.

Interestingly, a recent paper from Zhuo's lab found that rapid synaptic potentiation (changes in the strength of connections between different neurons) in the anterior cingulate mediates trace fear conditioning in response to noxious stimulation (Descalzi et al. 2012). When changes to the neuronal connection strengths in the cingulate were prevented, fear conditioning no longer occurred in response to noxious stimulation. Moreover, Zhuo and colleagues have found that noxious stimulation produces synaptic changes in neurons in specific cortical layers of the anterior cingulate cortex (Wei and Zhuo 2006) that mirrors the changes evoked by direct electrical stimulation of brain slices of the cingulate (Xu et al. 2008). Other researchers have shown that inhibiting neural activity in the cingulate can prevent conditioned place aversion, the preferred measure of affective pain in nonhumans, and also that facilitating neural activity in the cingulate can induce conditioned place aversion, even in the absence of noxious stimulation (Johanson and Fields 2004). Thus, one of the measures suggested to be among the best indicators of the affective component of pain in other mammals appears to be linked to activity in a very specific set of neurons.

This raises the fascinating possibility that we are entering a phase where understanding the basic operation of the affective component of pain at the level of synapses and individual neurons could be achieved in the relatively near future. And this, as I suggested above, might be seen as a prerequisite to being able to fully resolve some of the essential questions surrounding the unpleasantness of pain. Does unpleasantness always result in a change in the strength of the neuronal connections that underlie negative associations? Does it result in neurons firing that are inextricably part of the motivational system? Answering these questions could help tremendously with learning the true nature of pain's unpleasantness.

Again, however, we are held back by a lack of a full understanding of the role of pain's unpleasantness in humans. Imagine that researchers are successful

in pinpointing exactly which neurons are required for escape behaviors and trace conditioning in response to noxious stimulation in mice. Would this give us important information about the unpleasantness of pain in humans? Only insofar as the behaviors of the mice are reliable indicators of the same processes that generate the affective dimension of pain in humans. As it turns out, there are some indications that the cingulate is involved in aversive conditioning in humans. Milad et al. (2007) found that the same area of the anterior cingulate targeted for relieving pain in surgery was selectively activated by a conditioned fear stimulus associated with mild electric shock. Moreover, they found that the cortical thickness of the cingulate was correlated with a measure of fear response during the conditioning task. Relatedly, in another study, after humans were conditioned by being exposed to two tones, one paired with a loud unpleasant noise and one with a nonaversive noise, cingulate and insula activation were found to correlate selectively with the tone paired with the aversive condition (Buchel et al. 1999). So there is at least some evidence that trace fear conditioning in humans is linked to activity in the anterior cingulate cortex. But, in general, the precise role of the cingulate and insula during conditioning in humans is not yet understood.

As noted above, one research strategy that could offer some solutions but has been neglected is the study of patients with cingulate and insula lesions who have reported feeling pains but no longer finding them unpleasant. Are these patients also impaired on fear conditioning tasks? Are there other behavioral impairments present that could be used as indicators for the affective dimension of pain in other species? If it could be shown that lesions in humans' affective pathways led to both affectless pains and impairments on trace conditioning tasks, this would suggest that the studies on mice and rats locating mechanisms for escape behaviors and avoidance down to the level of specific layers of cortex in precise brain regions could tell us a great deal about the nature of the unpleasantness of pain.

To take a different example, Shackman et al. (2011) note that the same areas of the anterior cingulate cortex involved in pain have direct connections to the primary motor, premotor, and supplementary motor cortices involved in initiating movements as well as to the lateral periaqueductal grey, which is closely associated with fight or flight reactions. This suggests that activity in the anterior cingulate might be playing a role in the motivating behavior associated with pain. Experiments could be designed to see if humans lacking the affective component of pain also reliably lost the motivation to escape the pain.⁸

These examples could have important implications for the philosophical questions discussed above. First, detailed knowledge about the neural activity associated with the unpleasantness of pain would give us better criteria for attempting to assess if noncommunicative beings are capable of experiencing the unpleasantness of pain. Second, if it was found that the same neural circuits seem to be necessary for the unpleasantness of pain and for motivational signals to escape, this would tip the scales in favor of certain views in the philosophy of mind and axiology debates discussed above. In particular, it would lend credibility to both

the view that the *badness* of pain is inextricably connected to a desire to escape, and to the view that the phenomenology of pains needs to include some reference to motivation. In this way, pain research could be extremely important for resolving a number of important philosophical questions, but it currently has not progressed far enough to fill this role.

5 Conclusion

Understanding pain is important for a number of philosophical projects. And the unpleasantness of pain is of particular importance. Though the science of pain has helped us to learn quite a bit about the unpleasant dimension of pain, we still lack crucial information needed to provide complete answers to questions in value theory, applied ethics, and the philosophy of mind.

Given these limitations, the most important role philosophers can play is to collaborate with scientists to help nudge research towards answering some of these questions. Neither a pain science devoted exclusively to chasing clinically relevant results nor philosophers working primarily from intuitions can answer the most important questions about pain on their own. It is only through a collaboration that progress on the most important issues surrounding the unpleasantness of pain will occur.⁹

Notes

- 1 Melzack and Casey (1968) actually describe three determinants of pain behavior: “sensory” or “sensory-discriminative,” “motivational” or “motivational-affective,” and “central control” or “cognitive.” However, on my reading of their paper, the cognitive contributions to behavior don’t seem to be afforded the same ontological status as the sensory and motivational components of pain, which are both described as types of experience (p. 432). Rather, cognitive influences on pain behavior operate by modulating the other components: “Cognitive functions, then, are able to act selectively on sensory processing or motivational mechanism” (p. 433). Throughout this paper, I use “sensory” synonymously with “sensory-discriminative,” but I vary between “affective” and “affective-motivation” because I want to leave open the possibility that affect can be dissociated from motivational drives.
- 2 Korsgaard (1996), for example, writes, “the painfulness of pain consists in the fact that these are sensations which we are inclined to fight.”
- 3 It’s also worth noting that some theorists deny that the affective component of pain has its own phenomenology. For example, one might claim that the pain sensation gives rise to a further desire, and that the desire itself is unpleasant.
- 4 Parfit, however, when asked about this point during a lecture at Rutgers, referred to research showing that the unpleasantness of pain could be manipulated independently of the intensity of pains.
- 5 Since we generally desire to avoid things we dislike and generally dislike the things we desire to avoid, it is sometimes difficult to see the difference between the desire and the dislike views. Parfit (2011) offers a helpful contrast between the two views: “What we dislike is some sensation. What we want [desire] is not to be having a sensation that we dislike. Our desire could be fulfilled either by our ceasing to have this sensation, or by our continuing to have it but ceasing to dislike it. No such claims apply to dislikes, which, unlike desires, cannot be fulfilled or unfulfilled” (p. 54).

- 6 For an alternative perspective, see McCullagh (1997).
- 7 Subjects were asked to imagine themselves playing tennis or walking through their houses. The brain areas that were active in some of the patients matched those of healthy controls engaging in those imaginative activities.
- 8 One way of studying this would be measuring how hard people are willing to work to decrease the intensity of a pain.
- 9 Thanks to David Bain, Michael Brady, Jennifer Corns, and audiences at the University of Glasgow, York University, the University of Michigan-Flint, and the University of Pennsylvania for helpful comments on this essay.

8 When is a pain not a pain?

The challenge of disorders of consciousness

Valerie Gray Hardcastle

Introduction

I begin by stating what I take to be obvious (though others disagree): Scientifically speaking, we do not know what consciousness is exactly. In particular, we do not have a good definition for consciousness, we do not know what the psychological attributes relevant to conscious experience are exactly, and we have no precise notion what the neural correlates of consciousness are either. We have no good tests for consciousness, and we do not know how to identify conscious awareness in the brain. We do not understand the relationship between alertness and awareness, if there is one, nor do we understand the connection between cognitive processing and consciousness, if there is one. At best, we can point to some things that some people believe index some aspects of consciousness. But by the standards of contemporary science and medicine, that is not pointing to very much at all.

At the same time, there appears to be widespread agreement among philosophers (cf., Aydede 2005) and health care practitioners that pain is a conscious state of some type or other. The most widely accepted definition of pain was developed and later updated by a Taxonomy Task Force of the International Association for the Study of Pain (IASP). It considers pain a type of experience: “Pain is an unpleasant sensory and emotional experience that is associated with actual or potential tissue damage or described in such terms” (IASP 2012).¹ A key feature of this definition is that it “avoids tying pain to the stimulus. [Pain] is always a psychological state, even though we may well appreciate that pain most often has a proximate physical cause” (IASP 2012).

Such a definition can create challenges. If we do not understand what an experience – a moment or series of moments of conscious awareness – is exactly, then we also cannot readily understand pain, a specific type of experience. Neurologists are starting to argue, however, that new fMRI studies of minimally conscious and vegetative patients give us important clues into the neural circuitry required for consciousness. At the same time, other fMRI studies highlight areas of the brain activated by noxious stimuli in these patients. Some are putting these two threads of research together to conclude that vegetative patients cannot experience pain, but minimally conscious patients probably do (cf. Boly et al. 2005). In this essay, I shall suggest that both strands of argument are faulty.

In particular, I claim that understanding conscious experiences is not directly relevant for determining whether someone is in pain, and the type of disorder of consciousness one has is relevant neither to being in pain nor to how pain should be treated.

I'll proceed by outlining what the so-called disorders of consciousness are and why recent fMRI studies of some patients with disorders of consciousness are so important. I shall briefly discuss whether knowing someone is conscious is useful information to have for a potential pain patient. Finally, I shall turn to pain itself, looking at how we can detect pain using fMRI in normal subjects and brain-damaged patients, asking not whether these studies can tell us if a patient is *experiencing* pain, but whether this question is a useful one for medical personnel or caregivers to ask. As a preview of what is to come: it is not.

1 Disorders of consciousness

As medical treatments and interventions for patients with severe brain injuries have improved drastically over the past decades, we find many patients surviving significant damage who previously would have died. Many of these patients do not make a full, or even a partial, recovery. They exist in a kind of mental twilight, remaining in states known as “disorders of consciousness” (cf., Bernat 2006). These states lie on a continuum of increasing alertness and awareness,² starting with coma and ending with locked-in syndrome. In-between lie persistent vegetative states (PVS, also known as unresponsive wakefulness syndrome or UWS), minimally conscious states (MCS), and severe cognitive disabilities.

Not surprisingly, determining which patient is in which state can be exceedingly difficult, for most medical personnel are limited to personal interpretations of spontaneous behaviors. It also could be the case that calling these conditions “disorders of consciousness” is a misnomer. Often the patients appear to be unconscious, but it could be the case that they are merely unresponsive to their environments, and we are interpreting that as being unaware. For many doctors, and certainly for many loved ones, determining whether a patient is in fact still conscious is crucially important. More on this below.

A total lack of arousal characterizes coma. Comatose patients lie with their eyes closed, completely unresponsive to their environment and to any probing stimuli, though often basic reflexes remain (cf. Laureys et al. 2004). They exhibit no sleep/wake cycle, and most health care professionals believe that these patients are unconscious. Comas are normally caused by either a severe diffuse cortical or white matter injury or by bilateral lesions in the pontomesencephalic tegmentum or paramedian thalami in the brain stem (Owen 2008). Of the patients who survive, most emerge from coma in two to four weeks, though quite often they only move to a persistent vegetative state or a minimally conscious one.

Unlike patients in coma, the eyes of patients in a vegetative state will open and close in an apparent sleep/wake cycle. Because the patients appear awake, and they exhibit a wide range of reflexes, loved ones and family members often assume that the patients are conscious and can move purposefully. However, even

though these patients will move, the movements seem to be without purpose. Patients in a vegetative state cannot learn new responses, nor can they react to commands (Schnakers et al. 2006). In particular, the patients show no evidence whatsoever of sustained, reproducible, or voluntary movements in response to any sensory stimuli. While the cortex is often damaged in patients in a vegetative state, the thalamus most certainly is (Jennett 2002), though there is usually enough autonomic functioning remaining in the brainstem to allow the patients to survive if given nutrition and hydration.

“Minimally conscious” is a relatively new diagnosis for patients with severe cognitive deficits (Giacino et al. 2002). These patients show limited (but definite) evidence of sensitivity to their surroundings. Some patients move from coma to a vegetative state to being minimally conscious and then on to recovery; other patients can remain minimally conscious indefinitely. Just like patients in coma, minimally conscious patients exhibit reflexes, and just like patients in vegetative states, minimally conscious patients have apparent sleep/wake cycles and move without (obvious) purpose. But unlike patients in coma or in vegetative states, these patients can occasionally and inconsistently follow simple commands (“wiggle your fingers”), respond yes or no vocally or with gestures, move with purpose (e.g., pull away from a noxious stimulus), or speak a few words intelligibly. Adrian Owen (2008) hypothesizes that because the symptomatology of vegetative and minimally conscious states is so similar, the brain injuries underlying them are likely similar as well. However, given that minimally conscious patients can respond more reliably to stimuli than vegetative patients, it is likely that they have less diffuse cortical damage.

If patients can respond consistently and appropriately to commands or can regularly engage in goal-directed movements, then while they may still have a severe or moderate cognitive deficit, they are more than minimally conscious. These patients can engage in functional communication with those around them, indicating, for example, that they are hungry or uncomfortable. Some vegetative or minimally conscious patients pass through having a cognitive deficit on their way to recovery – patients with a traumatic brain injury are the most likely to recover – but a few patients just seem to wake up after years of being in stasis (Giacino and Kalmar 1997).

Finally, there are patients who apparently have completely normal internal brain responses to external stimuli, but they are completely paralyzed and are unable to respond to any prompts. These patients are known as “locked-in” (Plum and Posner 1966). They are unable to speak or move, though sometimes they do have control over their eyes or their eyelids. This condition occurs when the pons is lesioned such that all the descending motor pathways are interrupted. This leaves cognition virtually unaffected, but just about all behavioral responses are shut down.

At least some of those who investigate persistent vegetative state or minimally conscious patients believe that they can detect conscious awareness in non-communicative patients by looking at whether and how their brains appear to process cognitive information. Adrian Owen (2008) describes one way of

building up to this conclusion by following a method of measuring a patient's brain responses to an increasingly complex series of stimuli. For if a patient cannot hear, then that patient would not be able to interpret sounds. And if the patient cannot interpret sounds, then the patient would not be able to recognize individual words. And so on. So, if a patient's brain responds to acoustic signals in a near normal sort of way, by exhibiting a brain-stem auditory evoked potential on an EEG recording, for example, then one can see if the brain would then be able to recognize speech sounds as different from matched signal-correlated noise sounds. And if the patient's brain distinguishes between words and noise, then one can see whether the brain responds differentially to the patient's name or other overlearned and highly salient words.

Doctors have achieved some success with this method. For example, in a recent study, two out of seven patients diagnosed in a vegetative state showed brain patterns consistent with disambiguating complex sentences, what psychologists normally take to be the most difficult of verbal comprehension tasks (Coleman et al. 2007). These results, and others similar to it (cf., Owen et al. 2005), strongly suggest that some patients with purportedly the severest levels of brain damage can comprehend meaningful word strings. The question that immediately arises is whether this type of complex brain process indicates that the patient is conscious. If the patient exhibits relatively normal brain responses to stimuli that we normally understand that we need to be conscious to perceive and comprehend, does this mean that the patient is conscious as well?³

Adrian Owen believes he has found a way to answer this question. The next step in the hierarchy of tests to give patients with disorders of consciousness is to determine whether they can follow instructions to perform specific mental tasks. Can they imagine on command? He has discovered a few patients that seem to do so (cf., Owen et al. 2006). If he asks these patients to imagine playing tennis, then the supplemental motor areas show increased activity, just as they would were a normal subject given the same command. And if these patients are instructed to imagine walking through the rooms of their homes, their parahippocampal gyrus, lateral premotor cortex, and posterior parietal cortex become active, just as occurs in normal controls. Moreover, it seems fairly clear that these patients are indeed following instructions, for one does not see the same responses if the patients are told similar, non-instructive sentences, like, "The man enjoys playing tennis," or, "Shirley walked through the rooms of her house."

Adrian Owen and his colleagues are quite definitive in their conclusions, asserting that the fMRI patterns they saw "confirmed beyond any doubt that [their patient] was consciously aware of herself and her surroundings" (Owen et al. 2006, p. 1402). In addition, Martin Monti and Audrey Vanhaudenhuyse believe that functional MRI data can provide "clear evidence" that a patient is "aware and able to communicate" (Monti et al. 2010, p. 576). Members of Steven Laurey's Coma Science Group infer that "coactivation of specialized sensory cortices and frontoparietal areas [areas active in patients who can perform the imaginative tasks] seems both necessary and sufficient to generate conscious perception" (Boly et al. 2008, p. 1018). On the other hand, if "primary cortical

activation ... appears to be isolated from higher-order associative cortical activity [e.g., the supplemental motor areas],” then the patient is in a persistent vegetative state or has unresponsive wakefulness syndrome and is not conscious (Schnakers et al. 2012, p. 438).

Given the relative novelty and the relative paucity of these data, it is not a stretch to claim that such conclusions and assertions are premature, especially considering we really do not even know what we are talking about when we assert that someone (or something) is conscious. But more importantly, I shall be claiming that, relative to questions of these patients experiencing pain, these conclusions and assertions miss the mark entirely. It does not matter in the least what consciousness is or who has it. At least, it does not matter for this discussion regarding how to understand pain processing in patients with disorders of consciousness. Allow me to explain.

2 Consciousness and moral significance

Why is it that we should care whether others are conscious? Why is discerning a patient’s level of awareness important to doctors and the patient’s family? One reason we find knowing about the consciousness of others so compelling is that we assign moral worth to creatures based on their putative level of consciousness.

This connection abounds, for example, in moral arguments for vegetarianism (see, e.g., Singer 1975). After searching for the neural underpinnings of consciousness for most of his professional career, Christoph Koch (2012) has publicly converted to being a vegetarian because he sees too many similarities between human brains, and therefore human consciousness, and those of other animals. He argues thusly: Other animals are likely conscious, because they resemble us. Therefore, they likely feel pleasure and pain, as well as happiness, sadness, boredom, fear, and so on. Therefore, we should not eat them, just as we should not eat any sentient creature.

Logically speaking, how can Koch get to his final conclusion? He can if he has as a suppressed premise that conscious creatures have inherent moral worth in virtue of their consciousness.⁴ Koch appears to believe that conscious creatures have a special place in the moral hierarchy of our universe: they have intrinsic value. And because of this, we should treat them differently – we should treat them with care and respect. In particular, we should avoid treating them as a mere means to an end (like nourishing us as a meal).

Closer to our concerns in this essay, similar considerations arise in discussions around brain death and abortion. The United Kingdom uses a brainstem definition of death (rather than the cessation of activity in essentially the whole brain, as the United States does) because, it is argued, the loss of activity in the brainstem necessarily means the permanent loss of consciousness due to the loss of the reticular activating system, among other areas (cf., Kahane and Savulescu 2009). In the United States, the advisability of late stage abortions is undergoing renewed scrutiny, largely because questions over whether fetuses are aware of and

can experience pain, and, if so, at what point such experiencing begins, remain unanswered (or at least unanswered to everyone's satisfaction).

I am not going to argue against the idea that we should hold being conscious fundamental to imbuing a creature with inherent value; I am just accepting as given that this perspective is widespread in contemporary Western society. Many people (perhaps most people) believe that being conscious elevates one's moral status in the world. My only point here is that this is, in fact, a deeply held premise in most folks' moral reasoning – so deep, in fact, that we do not bother to defend that premise in moral discussions, science policy briefs, or medical texts. Indeed, most of us rarely even acknowledge its existence.

We can see the way these ideas are fundamentally woven into our cultural fabric through the case of Terri Schiavo. The Florida court battle over whether to allow her husband to remove her feeding tube depended on, among other things, whether Schiavo was in a persistent vegetative state or whether she was minimally conscious. The assumption all around was that if she were minimally conscious, if she had at least some fleeting awareness of her surroundings and the things in them, then her life should be preserved (see Horne 2009, Kahane and Savulescu 2009, McCullagh 2004 for discussion). The unargued assumption was that, other things being equal, her life was more valuable if she were on the right side of the conscious/unconscious divide.⁵

As an interesting side note, Florida law is actually written so that a surrogate decision-maker for a minimally conscious patient could in theory refuse artificial nutrition or hydration. Its statutes refer to the “absence of voluntary action or cognitive behavior of any kind” and the “inability to communicate or interact purposefully with the environment.” In addition, it authorizes surrogate decision-making for an “end-stage condition,” which it defines as “an irreversible condition that is caused by injury, disease, or illness which has resulted in progressively severe and permanent deterioration” (cf., Eisenberg 2008). This language does not exclude patients with diagnoses of minimally conscious states from falling under its statutory authorization, including perhaps Terri Schiavo. But I digress.

What I am going to focus on is the flip side of the deep assumption that being conscious carries moral weight. What we need to examine is whether being conscious is the *only* way to elevate our moral status. Could entities that are not conscious also have inherent value? The founder of the categorical imperative himself might have thought so, for rationality was the only requirement Immanuel Kant advanced for being an end-in-itself (1785). According to him, being able to choose a course of action to achieve a set of goals, as opposed to merely following instincts and desires, is what sets us humans apart from other animals. And it is this ability that gives us special moral status.

Others connect moral significance with the possession of interests. For example, Kahane and Savulescu write that, “it is a truism that interests matter morally” (2009, p. 11). What counts as an interest can vary widely, depending on what sort of ethics you happen to hold dear. Hedonists tie interests to being happy or experiencing pleasure. Deontologists hold that certain activities are intrinsically good, like developing talents or pursuing knowledge. Desire-satisfaction theorists

believe that satisfying our desires is what is important. And so on. For the purposes of this essay, it is not relevant whether any of these views are correct. If one could be happy, engage in certain activities, or have one's desires satisfied without being conscious, then we can conclude that there are ways of understanding moral worth that do not require being conscious.

Not everyone believes that we could do or have these things without being conscious. (Hassoun and Kriegel (2008) argue that to be a morally significant being requires consciousness of self.) The problem with these sorts of attitude, however, is that they separate conscious processing from unconscious processing. They imply that conscious cognition, perception, emotion, and thought, are important, but unconscious processes, like subliminal perception, cognition, affect, and so forth, apparently have no relevance for determining worth. These sorts of attitudes also assume that consciousness is separate and separable from our unconscious states. I submit that these positions are both mistakes.

We have known since the late 1800s that our brains process stimuli presented below the threshold of conscious awareness because subjects are able to respond to questions regarding the putatively unperceived stimuli above the level of chance (e.g., Sidis 1898). Classic studies conducted in the 1970s by Anthony Marcel (1974, 1983) further demonstrate that unconsciously or subliminally perceived stimuli influence decisions regarding consciously perceived stimuli. Investigations of memories of unconsciously perceived events, such as perceptions of things that occur while a patient is under anesthesia, provide further evidence of subliminal processing (cf., Merikle and Daneman 1996). Additional evidence for the existence of unconscious processing comes from the study of neurological disorders in which conscious perception of a variety of input is specifically impaired, while the ability to process and make use of such input remains preserved (e.g., blindsight, deaf hearing, numb touch, and the various agnosias) (Weiskrantz 1996, Moscovitch et al. 1997, Cowey 2001, Barton 2003). Many subsequent studies have explored the processing power of unconscious cognition (see Merikle and Daneman 1998 and Merikle et al. 2001 for review and discussion). We know that many stimuli are processed in quite sophisticated and deep ways without ever making it into consciousness. As John Bargh notes, the unconscious is not a "shadow" of our conscious minds. Instead, "the unconscious is not identifiably less flexible, complex, controlling, deliberate, or action-oriented than is its counterpart" (Bargh and Morsella 2008).

The point is that many important and sophisticated things appear to happen in our brains outside of conscious awareness. Many of these important and sophisticated things appear to be identical to what our brains can do consciously, too. Currently, we really have no idea what distinguishes conscious from unconscious phenomena in the brain. (I recognize that, given this fact, it is hard to maintain that the above outlined experiments do indeed test unconscious processing, as opposed to some other distinction that we think is allied with the conscious/unconscious divide. But we do not need this interpretation of the experiments for my main point, which is that we have a hard time delineating the difference between the two.) Consequently, it is likely a mistake to make too much

of the distinction between the two, especially from a neurological perspective. In particular, it is likely a mistake, from my vantage point anyway, to make the division between consciousness and unconsciousness what cleaves morally worthy creatures from the unworthy, for that difference is just too shaky and poorly understood to be able to support such an important partition.

3 Pain processing in patients with disorders of consciousness

Doctors and families often want to know whether their patients are conscious because they believe the answer should inform how the patients are treated. The question for many health care professionals is whether a patient can feel any pain or is suffering. Their answer to this question determines how much or whether the patient receives analgesia or other sedating medications. Terri Schiavo, for example, did not receive any pain medication after her feeding tubes were disconnected because the courts determined that she was in a persistent vegetative state and, hence, unable to experience pain or discomfort. Indeed, over half of doctors surveyed in Europe believe that patients in a persistent vegetative or unresponsive wakefulness state do not experience pain (Dermertzi et al. 2009), and it is reasonable to expect them to act on these beliefs by not providing analgesic medication for these patients during care or during the dying process after withdrawal of artificial hydration and nutrition (for discussion, see Ahronheim and Gasner 1990, Fins 2006).

Paralleling the IASP definition, the Multi-Society Task Force on Persistent Vegetative States decreed that “pain and suffering refer to the unpleasant experiences that occur in response to stimulation of peripheral nociceptive receptors and their peripheral and central afferent pathways or that may emanate endogenously from the depths of human self-perception” (1994). In other words, from their perspective, pain too is a conscious state, one caused by nociception and suffering. Therefore, if one is not conscious, then one cannot be in pain.

Indeed, the Multi-Society Task Force on Persistent Vegetative States concluded that the pain behaviors which patients in persistent vegetative states often exhibit – grimacing, posturing, crying, even racing heartbeats and hormonal fluctuations – do not indicate actual pain or suffering “unless they are consistent, sustained, and definitive in nature” (1994).⁶ While nociceptive stimuli can cause unconscious behavioral responses, as well as other autonomic or endocrinological reflexes, the task force states that this can happen without having any experience of pain or suffering if the brain has lost its capacity for self-awareness.⁷

But is asking whether a patient is conscious the right question to ask in order to determine the level of pain processing or of palliative care? To answer this question, let us compare how normal brains process pain with how the brains of patients with disorders of consciousness process pain.

We know quite a lot (but not everything) about how pain is processed in the brain. In a nutshell, stimulating our peripheral nociceptors activates a dedicated spinothalamic tract that goes to the thalamus and the midbrain, which then transmits pain information up to the cortex. This initial processing is normally

referred to as “nociception” and is largely reflexive. Pain processing itself is more complex, as it assigns affective value, behavioral motivation, and cognitive meaning to the stimulus.

Historically, this type of processing has been divided into two basic components, an affective-motivational processing pathway and a sensory-discriminatory processing pathway. Its fundamental importance was underscored in the middle of last century, as neuroscientists discovered two different spinothalamic tracts involved in nociception, a medial pathway involved in processing affect and a lateral pathway involved in sensory encoding. Neuroscientists later tried to connect activity in the primary and secondary somatosensory cortices to the sensory – discriminative aspects of pain processing, and activity in the cingulate, insula, and prefrontal cortices to the cognitive and affective aspects of pain processing. (In addition to these ascending pathways, there is also a well-defined descending network involving the prefrontal cortex, cingulate, periaqueductal gray, and the posterior thalamus that modulates the ascending system. This descending pathway can alter the brain’s responses to nociceptive inputs, as well as behavioral and perceptual outputs such that we see different responses to the same stimuli, if the stimuli are given in different contexts.)

However, in the 50 years since the dual-processing stream for pain was hypothesized, scientists have not made much progress in uncovering the details of its operations. Indeed, as Vania Apkarian maintains, evidence for division remains “fairly weak” (2012, p. 6). He, along with others, advocates a different perspective on pain processing. His view is that pain processing is really a subset of our general motivational and reward system. The basic idea is that acute pain, which normally signals tissue damage, first motivates us to escape or avoid our current situation in order to minimize physical harm, and then, when the pain stops, rewards us with a sense of relief. Together, our two reactions – avoidance and relief – not only protect our bodies, but they also contribute to our being able to predict the costs and rewards of competing behavioral plans. That is, we can learn from previous pain events what to expect in potential future decisions. In particular, we can learn how various short-term avoidance strategies will be rewarded by later senses of relief, all of which can be factored into our decisions regarding which behavior to select under a variety of circumstances.

This perspective on pain processing entails that there is no brain circuit specific to pain processing (cf., Iannetti and Mouraux 2010). Pain processing is simply part of our on-going risk-reward calculations. Indeed, it might turn out that Melzack’s (1989) original idea for a non-specific neuromatrix, a widely distributed network of neurons that crosses many areas of the brain and underlies pain perception, might be correct after all. But regardless of which theoretical approach is the correct one, we do know that pain processing involves many, many different areas of the brain; it is a very complex, widely distributed, and massively parallel process. Imaging studies of healthy subjects experiencing pain bear out the existence of the complex network (cf., Peyron, Laurent, and Garcia-Larrea 2000). They show widespread activation across many cortical and subcortical areas, the details of which need not detain us here.

What is important for us, though, is the sort of brain damage one typically sees in persistent vegetative and minimally conscious patients, and how that damage relates to pain processing in the brain. As I suggested above, we find more large-scale dysfunctions in the most cortical areas in patients in a vegetative state than in minimally conscious patients (Laureys et al. 1999, Laureys et al. 2004). These brain disruptions are presumably correlated with the degree of clinical impairment these patients have, and hence, they might be markers for a patient's relative level of overall cognitive deficiency. Important for understanding pain processing in these patients, imaging studies have shown that connectivity within and between the cingulate (a brain region strongly associated with processing pain) and the prefrontal cortex (part of the purported pain neuromatrix) is more disrupted in vegetative patients than in minimally conscious patients (Cauda et al. 2009, Vanhaudenhuyse et al. 2010).

Using positron emission tomography, Steven Laureys has examined how the brain reacts to nociceptive stimuli in patients in a vegetative state as compared with healthy controls (Laureys et al. 2002). In all the patients, only the midbrain, thalamus, and primary somatosensory cortex were activated in response to electrical stimulation of the median nerve, which demonstrates at least intact nociceptive processing and perhaps some higher-level processing as well. Later studies have found not only these areas to be active in persistent vegetative patients during pain processing, but also secondary somatosensory cortex, the cingulate, and the posterior insula (Kassubek et al. 2003). However, in no cases were brain activation responses close to what one finds in healthy controls. Vegetative patients' cortical areas appear to be functionally disconnected from one another and from other subcortical areas, which would presumably lead to incomplete pain processing (Boly et al. 2005). In particular, recent work suggests that vegetative patients lack many of the feedback connections from the association cortices (Boly et al. 2011). Investigators take this evidence to imply that vegetative patients lack the ability to integrate information from different areas of the brain, which, incidentally, some hypothesize is crucial for conscious experience. They conclude that whatever pain processing is going on in the brain of a vegetative state patient, it is likely to be unconscious and quite unlike what is going on in normal people's brains, which would include their conscious experience of pain (cf., discussion in Schnakers et al. 2012).

Baseline integrative capacities in minimally conscious patients do appear to be better preserved than those in vegetative state patients (Vanhaudenhuyse 2010). Not surprisingly, we also see brain responses to nociceptive stimuli in minimally conscious patients that are very similar to healthy controls, with activation in the thalamus, primary and secondary somatosensory cortex, insula, and cingulate (Boly et al. 2008). Also not surprising is that investigators conclude that minimally conscious patients are likely to have the brain integrity necessary for conscious awareness of nociceptive stimuli:

the co-activation of specialized sensory cortices and frontoparietal areas seems both necessary and sufficient to generate conscious perception. ...

We interpret the brain activation and functional connectivity patterns seen in patients in MCS as likely to show conscious perception of noxious stimuli.

(Boly et al. 2008, p. 1018; cf., Schnakers et al. 2012)

But regardless of whatever the conscious states of minimally conscious patients may be, their neural responses to nociception and pain are virtually identical to normal subjects.

I do not wish to adjudicate how vegetative state or minimally conscious patients' neural systems do or do not respond to nociception and pain relative to healthy subjects. Nor am I interested particularly in whether one feels pain in any of these conditions. It is sufficient to note that vegetative state and minimally conscious patients alike process nociceptive information, and their brains both also respond to pain to some degree or other, claims I believe that all researchers in this area would agree with, and claims that we can make without knowing exactly what consciousness is, or, for that matter, exactly how pain is processed in the brain. These facts are enough to raise the question regarding how one should treat vegetative state or minimally conscious patients in pain.

We already know that, in the main, doctors and other health-care providers assume that patients in a vegetative state are not conscious and therefore they can neither feel pain nor suffer. Most also assume that minimally conscious patients have some awareness of their bodily conditions and therefore they likely feel pain and can suffer. Consequently, they feel no real impetus to provide pain relief to patients in vegetative states.

Even when we look at what the neurologists who study pain processing in patients with disorders of consciousness say about treating these patients, we see similar lines of reasoning. At best, the authors point out that upwards of 40% of patients in putative persistent vegetative states are misdiagnosed (cf., Schnakers et al. 2009), so they urge prudence in providing analgesics to vegetative state patients:

The data suggest that these MCS patients could perceive pain. On the contrary, VS patients showed a functionally disconnected brain activity, suggesting the absence of an integrated pain perception. Considering these results, adequate analgesic treatment has to be provided in MCS patients. The issue is much more complicated in VS patients. Given the high rate of misdiagnosis ..., if we decide not to administer analgesic treatment in the presence of a potential painful experience (e.g., contractures or fractures), there is a real probability for not treating a patient erroneously diagnosed and, hence, for not treating a patient who perceives pain.

(Schnakers et al. 2007)

Considering the levels of clinical uncertainty, pain treatment should be considered in all patients diagnosed as being in a VS/UWS or MCS.

(Schnakers et al. 2012, p. 442)

It is our opinion, based on both extensive clinical experience and recent data questioning the ability of the bedside exam, even when sophisticated, to delineate awareness in at least some patients with DOCs, that one is likely taking the safer course by treating all DOC patients as if they had the potential to perceive pain and suffer.

(Schnakers and Zasler 2007, p. 624)

I suggest that it is important to know that both persistent vegetative patients and patients in minimally conscious states process nociceptive stimuli and engage in what we would commonly recognize as at least some aspects of pain processing. I submit, however, that it is not important to know which aspects of the processing are conscious. Here is why.

4 The impact of pain

In a nutshell, processing pain in and of itself has a negative effect on the brain and body. We gain support for this conclusion through several very different strands of evidence. I will briefly outline three of them.

Preterm infants normally undergo multiple painful experiences – heel sticks, intravenous catheters, chest tubes, endotracheal suctioning, surgery. Indeed, the sickest of premature babies are subjected to an average of 750 procedures before their discharge (Porter et al. 1999). Premature infants in Canadian hospitals normally endure two to eight painful procedures each day. And yet, analgesics are provided in fewer than 10% of the cases (Johnston et al. 1997, Stevens et al. 2003). However, investigations of nociception in neonatal patients indicate that preterm infants are more sensitive to pain than full-term infants and, more important, that acute nociceptive stimuli lead to extended episodes of hyperalgesia, during which non-painful stimuli can induce chronic pain (cf., Anand 1998, Page 2004). In addition, nociceptive stimuli in neonates have been linked to early intraventricular hemorrhage and periventricular leukomalacia, a type of white-matter brain injury. Finally, administering analgesics to neonates is correlated with fewer episodes of these types of injuries (Anand 1998). In short, untreated pain in these infants seems itself to cause brain damage, regardless of whether premature infants consciously experience pain. (Most neonatal doctors believe they cannot; most parents of premature infants believe that they can.)

And we do not need to focus on preterm infants to see long-term effects of painful stimuli. Coded pain scores during immunization in male babies are linearly related to type of circumcision analgesia (no circumcision, topical agents, and placebo analgesics) (Taddio et al. 1997). There are also some data linking early pain experiences to permanent changes in spinal cord processing as well as a compromised immune system and even some behavioral disorders, such as hyper-vigilance, sleep disturbances, avoidance behaviors, and feeding problems (cf., Page 2004, Mitchell and Boss 2002). Regardless of the considerable discussion regarding whether young infants are conscious or, if conscious, whether they

can remember painful events, traces of those events linger in the body and the brain, altering their developmental trajectory in fundamental ways.

We can see similar permanent effects with chronic pain. Chronic pain is represented in different areas in the brain than acute pain, because, as it turns out, the brain reorganizes itself when it is in chronic pain. Most types of chronic pain shift away from what might be referred to as simple nociceptive processing to more affective reactions. Moreover, we see specific ongoing non-pain effects; for example, the baseline level of activity in the insula and anterior cingulate is much higher in chronic pain patients than in control subjects.

In addition, how the nucleus accumbens connects to the rest of the brain is different for healthy subjects than chronic pain patients. In normal subjects, the nucleus accumbens and the insula are tightly connected to each other. But in chronic pain patients, the nucleus accumbens shifts its functional connectivity to medial prefrontal cortex (Baliki et al. 2010). And the greater the intensity of chronic pain the patients feel, the stronger the correlation between activity in nucleus accumbens and medial prefrontal cortex. As a result of this rewiring, nucleus accumbens activity differs between healthy subjects and chronic pain patients for instances of acute pain as well. Finally, this change in brain connectivity is a functional rewriting that goes far beyond pain processing, for, in contrast to normal subjects, the brains of chronic pain patients show no significant response to either a monetary reward or a monetary loss (Apkarian 2012). In other words, chronic pain puts stress on our motivational and reward systems such that these systems become quite dysfunctional, sometimes to the point of not functioning at all. This change in functionality is so large that it distinguishes between normal subjects and chronic pain patients with an accuracy of more than 90% (Baliki et al. 2012).

These results are even more remarkable because they have been shown to be independent of subjective pain perception; that is, they are independent of one's reported experience of pain. The distorted brain responses are not reflected in the pain ratings given by chronic pain patients, and chronic pain patients report being unaware of any impact their pain sensations have on any aspect of their cognitive processes. Nevertheless, chronic pain significantly influences later actions and decisions.

Pain processing also changes brains morphologically. With long-term or repeated pain, we can measure a significant decrease in gray matter in the areas associated with the "neuromatrix" in the brain, especially in the anterior cingulate cortex, right insular cortex, dorsolateral prefrontal cortex, amygdala, and brainstem (cf., Rodriguez-Raecke et al. 2009). A few studies suggest that areas that exhibit the most change in gray matter density also include the hippocampus, multiple lateral frontal regions, and portions of the occipital lobe, suggesting that the morphometric changes are not limited to pain-specific regions (Baliki et al. 2011). Either way, the observed morphological differences in chronic pain conditions correlate to the length of time pain patients have been suffering as well as the intensity of their pain (Baliki et al. 2011).

All of these lines of evidence suggest a new way to think about treating pain. Regardless of level of consciousness in patients, it is clear that processing pain

information is bad for their brains and other functional systems. It is also clear that blocking brains from processing pain information prevents these pain-related disruptions. Therefore, it seems clear (to me at least) that whether a patient is conscious has little bearing on whether one should treat that patient's pain. We should treat pain because not to do so is detrimental to the system. All else being equal, minimizing pain activation is always preferable.

The question of how and whether to treat putative pain in patients with disorders of consciousness is a real issue, too, for there are a whole host of reasons why such a patient might be processing pain information. In the acute stage of injury, patients often experience fractures, internal injuries, soft tissue injuries, and then the pain associated with invasive surgical procedures. In the chronic stage, patients can exhibit spasticity, contractures, pressure sores, ischemia, peripheral nerve injuries, and postsurgical incisional pain. In addition, headaches are very common after traumatic brain injury, as well as central pain from nerve damage or stroke. It makes a difference how we approach pain processing in patients with disorders of consciousness.

I conclude: How and whether to treat patients with severe disorders of consciousness should not depend on the type of disorder or level of consciousness, alertness, or awareness. It should instead depend on known brain circuitry and neurophysiological and behavioral responses. We should focus on maximizing the good for the system as a whole, regardless of consciousness. Therefore, we should treat pain in all patients, even those we believe are not conscious.

Philosophers tend to focus on the conscious aspects of pain in our bodily systems. However, regardless of consciousness and whatever perceived "badness" of pain due to the conscious awareness of it, pain is also detrimental because it can permanently damage our neural (and other) structures. This possibility for long-term damage, in itself, is a reason to use analgesics to diminish our brains' responses to pain information or to otherwise avoid pain if at all possible.

Notes

- 1 While this definition does not claim that pain is conscious explicitly, I do believe that this is an easy inference to make, given that pain is defined as an experience of emotion. Rare is it that people talk about unconscious emotional *experiences*. If one conceives of an emotion as something happening outside of awareness, then it is described in terms other than that of an experience.
- 2 These are being used as technical terms, essentially referring to one's orientation and ability to react to one's environment.
- 3 This question, of course, begs the question regarding what is needed, either neurophysiologically or psychologically, for consciousness. This issue does not seem to bother clinicians or researchers who are pursuing these studies. It does, however, bother me.
- 4 Michael Brady has pointed out that one could make a utilitarian argument to the same end. He is quite right.
- 5 The courts ruled that Schiavo was in fact in a persistent vegetative state, that her husband had the legal right to determine the course of her medical treatment, and that the legislative and executive branches of the government were blocked from intervening by the establishment clause of the First Amendment. She died 13 days after her feeding

tube was removed, and autopsy revealed that her brain had atrophied significantly in the 15 intervening years after the hypoxic event.

- 6 Sustained grimaces of course are not definitive of consciousness, but they are one measure we often use to conclude that a conspecific is in pain.
- 7 The task force has been criticized for making the bar for experiencing pain so high. Intuitively, many people believe that a creature can experience pain without being aware of a self – dogs and infants might be two examples of this. The task force respectfully disagrees but provides no real data to support this high bar.

9 The first-person in pain

Frédérique de Vignemont

Introduction

‘Are you in pain, dear mother?’ ‘I think there’s a pain somewhere in the room,’ said Mrs Gradgrind, ‘but I couldn’t positively say that I have got it.’
(Dickens, 1854, p. 234)

The very idea of feeling a pain in a limb which does not seem to be ours is difficult to frame, perhaps unintelligible.
(Dokic, 2003, p. 325)

When in pain, I experience disorder as located within a boundary of whose exterior I can have no similar awareness; I am aware that a part of my body – not just of that body – is disordered.
(Bain, 2003, p. 523)

Human’s experience of pain is strictly dependent from the way we represent the body itself and from the sense that it is *my* body that is undergoing a certain experience (i.e., body ownership).
(Pia et al., 2013)

Mrs Gradgrind’s reply strikes us as hard to understand, or even conceive. Is she claiming that she does not take herself to be the person feeling the pain? Or is she describing that she does not take the pain to be located in her own body? It is hard to say, but for the sake of this chapter, I shall assume the latter interpretation: Mrs Gradgrind feels pain in a body that does not seem to her as being her own. This is precisely what Jérôme Dokic claims to be ‘unintelligible’ (see also Martin, 1995, p. 270), and indeed, leaving aside empathy for pain, we intuitively assume, like David Bain and Lorenzo Pia and his collaborators, that we feel pain in our own body only. But why is it so? And is it always true?

Here, it is important to distinguish two distinct claims. The first is simply a physiological claim. It is a biological fact that in our world the nociceptive system is such that it is connected only to our own body and to no other bodies. But Dokic’s and Bain’s claim is different. It does not concern which body we can feel pain for, but rather how we *experience* the body for which we can feel pain. What is at stake here is the relation between pain and what is known as the *sense*

of *bodily ownership*; that is, the awareness of one's body as one's own. According to Bain and others, one cannot have pain without the sense of bodily ownership. This view can be phrased as follows:

In-my-body view: If one feels pain in *x*, then one is aware of *x* as being part of one's own body.

In this chapter, I shall explore this view in more detail. I will first analyse the reasons for which one may argue that pain is inseparable from the sense of bodily ownership. It may be because the sense of ownership is given by pain (Martin, 1995), or it may be because pain requires the sense of bodily ownership (Pia et al., 2013). I will then assess whether it is really impossible to feel pain in an 'alien' body in light of borderline cases of ownership. I will conclude by comparing our experience of pain with our experience of threats in their relation to the sense of bodily ownership.

1 From pain to the sense of ownership

How am I aware that this is my own body rather than someone else's? A good starting point for analysing the sense of bodily ownership is to compare the awareness of my own body with the awareness of other bodies. Through vision, audition, touch, smell and taste, I have access to bodily properties, whether they are instantiated by my body or by other bodies (e.g. skin colour, bodily posture). Clearly, this type of bodily awareness, which is from the outside, does not suffice for the sense of bodily ownership because it gives also access to other people's bodies. Hence, one can relatively easily fail to recognise the body that one sees in a photograph or in a mirror as being one's own. As Brewer says, the visual body does not bear the 'stamp of ownership' (1995, p. 305).

Unlike other bodies, not only do I perceive my body through external senses, but I also receive information through what I call body senses (including touch, proprioception, interoception, nociception, and the vestibular system), which give rise to bodily experiences. The classic five senses do not exhaust the list of ways of gaining information about my own body. Let us imagine that I wake up in the middle of the night. I feel my heart beating too fast. My left arm feels in an awkward posture, which is painful. It feels nice to stretch on the bed and to feel the contact of the cold sheet on my skin. I feel too warm. I am thirsty. I get up, but I lose my balance. These various types of information about my body are not directly available about other people's bodies.¹ One may then suggest that there is no need to go further in the analysis of the sense of bodily ownership. Indeed one has never bodily experiences for any body other than one's own. Thanks to their privileged relation to one's body, bodily sensations may thus be all it takes to be aware of one's body as one's own (Brewer, 1995; Cassam, 1997; Dokic, 2003; Martin, 1995). For instance, Martin (1995) defends the view that the sense of bodily ownership is not an additional quality to the sensory qualities of bodily experiences and that it is 'somehow already inherent within them' (p. 278).

On his view, it is sufficient to feel pain (or any other bodily experience) in a body part to be aware of this body part as being one's own. Pain can thus rationally ground one's judgement of ownership: if I feel pain in *x*, I am entitled to judge that *x* is part of my own body.

2 From the sense of ownership to pain

Theories of the sense of bodily ownership conceive of pain as being one route – among others – to the awareness of one's body as one's own. This conception assumes no difference between pain and other types of bodily experiences. According to Martin (1995, p. 277), the sense of ownership is 'possessed by all located sensations'. On that respect, there is nothing special about pain compared with other bodily sensations.² However, we shall now see that pain has some distinctive features, which may reinforce its relation to the sense of bodily ownership (see Vignemont, 2017a for a systematic comparison between pain and touch). For some, the sense of bodily ownership is even a necessary requisite for pain. This may be so for two reasons, first because of the spatial peculiarities of pain, and second because of its motivational role.

2.1 *The location of pain*

There has been extensive discussion about the spatial ascription of pain in the literature (e.g. Noordhof, 2001; Tye, 2002). Typically, it has been asked how we should interpret the term 'in' when one reports feeling pain in one's hand, since such ascriptions do not seem to follow the rule of spatial transitivity. For example, I feel pain in my thumb and my thumb is in my mouth, but I do not feel pain in my mouth (Block, 1983). Here I will not engage in this debate and rather focus on another interesting feature of the spatiality of pain, which has been more neglected. In brief, the location of pain seems to be more constrained than the location of tactile sensations.

It is commonly accepted that tactile sensations can be felt beyond the biological boundaries of one's body, at the tip of a tool for instance (Gibson, 1979; Martin, 1993; O'Shaughnessy, 1980; Vesey, 1961). One might say, for instance, that the blind man is aware of the obstacles on the floor *at the end* of his white cane rather than on his palm.³ However, there is no comparable sense in which the blind man could be said to feel pain at the end of his cane. Of course, he can wince when the cane falls on the floor. But he does not feel pain as being localised in the cane. The fact is that one uses tools in harmful situations in which one would not use one's limbs. This is so because if a tool is damaged, one does not feel hurt. Hence, although there is a sense in which one can be said to feel resistance or vibration in a tool, the same cannot be said of pain. This is not to say that one can feel pain only within the limits of the biological body. Unfortunately, pain in phantom limbs does exist and is extremely unpleasant. So why can one feel pain in phantom limbs but not in tools?

One notable difference between tools and phantom limbs is that the amputee feels the phantom limb as part of her own body whereas the blind man does not feel the cane as part of his body. One way to interpret the spatial difference between pain and touch is then to claim that pain can be localised only in the body that one experiences as being one's own. On this view, the sense of bodily ownership is a prerequisite for experiencing pain in a body part. One needs to be aware of *x* as being one's own body part to be able to localise pain in it.

2.2 Bodily care

Let us now consider another difference between pain and touch. Most of the time tactile sensations are devoid of affective valence: I feel the pressure of the table on my arm, but unless it feels uncomfortable, my tactile sensation is affectively neutral and has no immediate implication for action. By contrast, along with itches and tickles, painful sensations have affective and motivational dimensions: they feel unpleasant and they motivate me to act accordingly. Whether one defends an evaluativist or an imperativist view of pain, it has been recently argued that pain plays a motivational role because one cares about one's body (Bain, 2014; Klein, 2015a). In brief, one protects one's body from injury because one cares about it. On this view, bodily care is a necessary condition for the motivational role of pain.

[A] pain will represent damaged states as bad [which is what its unpleasantness and motivational force consists in] only to a subject who cares about her own body.

(Bain, 2014, p. 315)

Let us consider in detail the imperativist view. According to Klein (2015b), pains consist in bodily commands. However, as he notes, we do not obey all commands. We do so only when we accept the source of the command as having the right authority: acknowledging the practical authority of the source suffices to motivate and to give us reason to act. Klein then argues that in the case of pain the body constitutes our practical authority. By recognizing the authority of the body, we are bound to at least consider the actions that will keep us alive:

Of course, we accept our bodies as authorities for good reason. Our body is important to us. We care about it. We are in bad shape if it doesn't work. We cease to exist when it does.

(Klein, 2015b, p. 81)

The difficulty for Bain's and Klein's views is that there can be many bodies one can care about, while pain seems to play a motivational role for one's own body only. Their notion of bodily care therefore must include that one has singled out the body that one must care about, namely, one's own body. In a nutshell, it presupposes a sense of bodily ownership (Vignemont, 2015). If we were to

rephrase Klein's imperativist view, for instance, we would claim that it is because I experience the body that commands me as my own that I obey it.⁴ Consequently, one should conclude that the sense of bodily ownership is a prerequisite for the motivational role of pain: one needs to be aware of *x* as being one's own body part for one's pain to motivate protective behaviour.

3 Alien pains

We now have a better understanding of why it does not seem intelligible to feel pain in a body part that does not seem to be ours. It is now time to assess whether the in-my-body view is empirically confirmed. To do so, I will examine the following claims:

- (a) If there is a sense of ownership, then there can be pain.
- (b) If there is no the sense of ownership, then there cannot be pain.
- (c) If there is no pain, then there cannot be a sense of bodily ownership.

With the help of various borderline cases of ownership and disownership, I will show that pain and the sense of bodily ownership are separable.

3.1 *Pain in an alien body*

Does it suffice for *x* to feel as being one's own for one to be able to feel pain in it, no matter whether *x* actually belongs to one's body or not? To answer this question, we may consider the Rubber Hand Illusion (Botvinick and Cohen, 1998). In the classic set-up, participants look at a left rubber hand presented in front of them, while their own left hand is hidden behind a screen. The experimenter then simultaneously strokes with two paintbrushes both the participant's hand and the rubber hand, either synchronously or asynchronously. After a couple of minutes, the experimenter asks participants whether it seems as if the rubber hand were part of their body. Participants then reply positively, but only after the synchronous condition. They further describe feeling their tactile sensations to be located on the rubber hand. What about pain? We have seen that one does not experience referred pain in tools, but can one experience it in the rubber hand that seems to belong to one's body? Unfortunately, the results are not conclusive. A few studies on pain using this illusion claimed that participants could feel pain in the rubber hand, but these experiments suffered from various methodological issues. Either pain was not directly measured (Mohan et al., 2012) or it was badly measured (Capelari et al., 2009). Let us just note a study by Valenzuela-Moguillansky and her colleagues (2011): both the rubber hand and the subject's hand were synchronously stroked. Immediately after that, they both received a painful stimulation. Subjects were then asked whether they felt pain *in the rubber hand*. They only mildly agreed (rating at 4 on a scale from 0 to 10). Still, they agreed more than when they did not feel the rubber hand as being their own.

More convincing may be the case of patients with what is known as the embodiment delusion. After a lesion in the right hemisphere, these patients report that another person's left hand is their own (Pia et al., 2013). For instance, when they see in front of them both their own left arm and the experimenter's left arm, they claim that the latter belongs to them. I shall call it the embodied arm. When the embodied arm moves, they report feeling their arm moving. Furthermore, they reach for the embodied hand when asked to reach for their own left hand. When asked to name the colour of the object in front of their own hand, they name the colour of the object in front of the embodied hand. Finally, they vividly react when the embodied hand is threatened (Garbarini et al., 2014). Their delusion of embodiment, however, requires the other person's hand to be next to their own hand. When only their own hand is on the table, their performance is normal. Interestingly, these patients can sometimes report pain when the embodied left hand is injured. Pia and his colleagues (2013) compared pain intensity judgements after noxious stimulation either on the patient's hand or on the co-experimenter's hand. They found that the patients reported feeling pain with the same intensity when their own left hand was hurt and when the embodied left hand was hurt. It seems unlikely that this result merely reveals a kind of empathy, because this effect was found only when the embodied arm was hurt, and not when it was the other arm that was hurt. Hence, it was only when the other person's hand felt as their own that the patients reported feeling pain there.

This shows that one can feel pain beyond one's biological boundaries as long as one feels it in the body that one takes oneself to have. This was already known for phantom limbs. We now know that this can be true also for limbs that are physically real, including for limbs that are biological parts of another body. It thus seems true that sometimes, at least, where there is a sense of bodily ownership, there can be pain.

3.2 *Pain in the body that feels as alien*

In order to determine the relationship between the sense of bodily ownership and pain, one should not look only at cases of bodily ownership; one should also analyse what happens in the *absence* of the sense of bodily ownership, and more specifically, at disownership cases. One will then be able to determine whether the sense of ownership is a prerequisite for pain.

The rubber hand illusion may shed light on that question too. It has indeed been argued that when one feels the rubber hand as one's own, one no longer feels the biological hand (which is stroked in synchrony with the rubber hand) as one's own (Moseley et al., 2008; Tsakiris, 2010). Roughly speaking, one can have only one left hand, and the rubber hand replaces the biological hand. In favour of this view, it was found that the temperature of the biological hand decreased, and that participants were slower in making judgements about tactile stimuli presented on the biological hand. If the sense of ownership were a necessary condition for pain, and if the biological hand were disowned because of the rubber

hand, then one should no longer be able to feel pain in the biological hand. One might then expect pain processing to be disrupted, or at least, a modulation of pain by the illusion, something like a sensation of relief. The rubber hand illusion, however, is not a way to cure pain. Several studies found no decrease of pain intensity when noxious stimuli were applied to the biological hand (Kammers et al., 2011; Martini, 2016; Mohan et al., 2012; Valenzuela-Moguillansky et al., 2011).⁵

But do these results invalidate the in-my-body view? Not necessarily. There is an alternative interpretation, according to which there is simply no sense of disownership of the biological hand in this illusion. The fact is that the evidence in favour of disownership is relatively weak. In particular, the slowing down of tactile processing found by Moseley and colleagues (2008) can be explained by other factors than disownership. A similar decrease in tactile performance has indeed been found following prismatic displacement independently of any feeling of disownership (Folegatti et al., 2009). What these studies have in common is the fact that participants see their hand at a location different from where they feel it. This visuo-proprioceptive conflict is most probably responsible for the disruption of tactile perception in the rubber hand illusion, rather than putative disownership of the biological hand.

We thus need to look at clearer cases of disownership, and more specifically at patients who experience their hand as alien. If the sense of ownership were a necessary condition for pain, then they should not be able to feel pain in the hand that feels as alien. Consider, for instance, the neurological condition of somatoparaphrenia, which follows a lesion of the right parietal lobe. Not only do somatoparaphrenic patients feel their hand as alien, but they are absolutely convinced that it is not their own hand and that it belongs to another individual (Vallar and Ronchi, 2009). It is true that sometimes patients with somatoparaphrenia also suffer from anaesthesia, and their ‘alien’ body part feels numb, but this is not always true. In particular, some patients with somatoparaphrenia can report feeling painful sensations in the hand that they disowned. For example, it has been reported that a patient cried out of pain when the examiner pinched his ‘alien’ hand (Melzack, 1990). Another patient asked his doctor:

Patient: I still have the acute pain where the prosthesis is. Examiner: Which prosthesis? P: Don’t you see? This thing here (indicating his left arm). The doctors have attached this tool to my body in order to help me to move. But it’s completely useless and very painful (...) Once home could I ask my wife, from time to time, to remove this left arm and put it in the cupboard for a few hours in order to have some relief from pain?

(Maravita, 2008, p. 102)

The situation is even clearer in patients with xenomelia (also called Body Identity Integrity disorder). Patients with xenomelia have apparently normal sensory and motor functions. Yet they have an overwhelming desire to be amputated of one of their perfectly healthy limbs, and when surgeons agree to cut their undesired

limb off, they feel relieved. One reason they put forward is that the undesired limb does not feel to be part of their own body (First, 2004; Hilti et al., 2013). Yet they can still feel pain in the limb they want to cut off.

3.3 Threatening alien hands

We have just seen that patients with xenomelia apparently show no disruption of their nociceptive system and that they can experience pain in their legs, although it seems to them that their legs do not belong to their body to such an extent that they want to get rid of them. What is possibly even more surprising is how they react when their legs are under threat. One should expect them to react normally, since their nociceptive system is intact and threat detection and pain have a lot in common. Recent results in neuroimaging actually show that the same areas are partly activated when one is in pain and when one perceives a threat approaching one's body through vision or audition (Legrain et al., 2011). One may even claim that responses to threat are guided by the anticipation of forthcoming pain (Haggard et al., 2013). Yet patients with xenomelia do not react to threat, although they can feel pain (Romano, 2015). In one study, patients saw either a Q-tip or a syringe approaching either the limb, for which they wanted amputation, or the contralateral limb, for which they showed no abnormal attitude. The experimenter then measured their skin conductance response (SCR), which normally increases when seeing one's body under threat. As expected, when the syringe approached the normal limb, the SCR increased, but it increased significantly less when the syringe approached the unwanted limb. By contrast, when the painful stimulus is actually applied to the skin, the SCR increased significantly more when it is on the unwanted limb than on the contralateral limb. In brief, they react less to threats but more to pain for their 'alien' limb.⁶

The reverse pattern can be found in the RHI. On the one hand, we have seen that it is difficult to induce referred sensations of pain in the rubber hand. On the other hand, participants in the rubber hand illusion react vividly when the fake hand is under threat. This is so only when they report it as their own after synchronous stroking, and the strength of their reaction is correlated with their ownership rating in questionnaires (Ehrsson et al., 2007). Threat response has actually become the main measure of the rubber hand illusion. It is found not only at the physiological level (as measured by the increase in SCR), but also at the motor level. In one study, participants were in an immersive virtual reality system, which induced them to feel ownership over a virtual hand. The virtual hand was then attacked with a knife. Although participants were instructed not to move, their event-related brain potentials revealed that the more they felt the virtual hand as their own, the more the perception of threat induced a response in the motor cortex (Gonzalez-Franco et al., 2013). Roughly speaking, when one observes the rubber hand that one experiences as one's own being threatened, one automatically withdraws one's hand. It may then seem that the sense of bodily ownership is more tightly linked to threat reaction than to pain.

3.4 *The painless body*

We have seen the consequences (or more exactly the lack of them) of the absence of the sense of bodily ownership for pain. Let us now assess the consequences of the absence of pain for the sense of bodily ownership. Patients with congenital pain insensitivity are characterised by dramatic impairment of pain sensations since birth, caused by a hereditary neuropathy or channelopathy. They show a complete lack of discomfort, grimacing, or withdrawal reaction to prolonged pinpricks, strong pressure, soft tissue pinching, and noxious thermal stimuli. Interestingly, patients with congenital insensitivity to pain can describe feeling their body like an external object, a kind of tool. An 18-year-old boy with congenital pain insensitivity reported:

A body is like a car, it can be dented but it pops out again and can be fixed like a car. Someone can get in and use it but the body isn't you, you just inhabit it.

(Frances and Gale, 1975, pp. 116–117)

It thus seems that Descartes (1641) was right: 'Nature likewise teaches me by these sensations of pain, hunger, thirst, etc., that I am not only lodged in my body as a pilot in a vessel' (*Meditation VI*). Without pain expertise, one is only as a pilot in a vessel. The unity between the body and the self appears to be broken. This leads patients with congenital insensitivity to pain to a range of self-harming behaviours, including burns and auto-amputations of fingertips and tongues (Danziger and Willer, 2009; Nagasako et al., 2003). Similar self-inflicted injuries can also be found in animals raised in isolation preserved from pain (Melzack and Scott, 1957; Lichstein and Sackett, 1971). Injuring this specific body may then be a way to test whether it is one's own or not. Feeling something can then strengthen a fading feeling of confidence in the sense of bodily ownership, whereas feeling nothing can confirm that this body has nothing to do with one. Or it might be that one simply feels no sense of ownership whatsoever. If so, one is not voluntarily injuring oneself, one is rather injuring *a body*, which happens to be one's own. In other words, when patients bite their fingertips, they are not fully aware that they are biting their *own* fingertip. Failing to feel one's limb as one's own leads to a failure to protect it from harm.

4 *Who cares whose body is in pain?*

We have seen that the relationship between pain and the sense of bodily ownership is far more complex than predicted by the in-my-body view. In brief, one can experience pain to be located:

- in one's hand, which one experiences as being part of one's body;
- in one's hand, although one no longer experiences it as being part of one's body;
- in a phantom hand, which one experiences as being part of one's body;

- in another individual's hand, but only if one experiences it as being part of one's body;
- but not in a tool, which one does not experience as being part of one's body.

To make the story even more difficult to understand, we have seen that threat perception does not follow the same pattern. In brief, one reacts to threat for:

- one's hand, which one experiences as being part of one's body;
- a phantom hand, which one experiences as being part of one's body⁷;
- another individual's hand, but only if one experiences it as being part of one's body⁸;
- but not for one's hand when one no longer experiences it as being part of one's body;
- and not for a tool, which one does not experience as being part of one's body.

What is the difference between pain and threat perception? And why does it matter for the sense of bodily ownership? Finally, we have seen that pain deprivation alters the awareness of one's body as one's own, which in turn can result in self-inflicted injuries. Based on these findings, we can neither claim that pain and the sense of ownership are inseparable, nor that they are fully independent. But what is their relationship exactly? What role is played by the sense of ownership? And what role is played by pain?

4.1 Is the sense of bodily ownership necessary for feeling pain?

One way to better understand the dissociation between pain and the sense of bodily ownership is to consider the equally counterintuitive dissociation between reaction to pain and reaction to threat. It may seem that we react to pain and we react to threat for the same reason, namely to avoid (further) injury or (further) pain. One thus expects little difference between the two. However, we have seen that whereas feeling pain does not involve experiencing one's body as one's own, reacting to threats does. The question then is why there is such a difference. In order to answer this question, we need to track the other aspects on which pain and threat reaction differ. Unfortunately, the notion of threat has been less extensively discussed in the philosophical literature than the notion of pain. Threats can be conceived of as a specific type of relational property between the environment and the agent. According to Gibson (1979), they are negative affordances: they describe how the environment is arranged while informing how one should move within this environment to avoid them. Threats are subject-dependent (what is threatening for me can be not threatening for another individual), and space-dependent (an event is more or less threatening depending on its location). For instance, I hear the bee close to my face as threatening, but not the bee on the other corner of the room next to you. Something is thus an external threat only if it stands in an appropriate spatial relation to one's own body. Since threats occur in external space, a space in which there are many

bodies, one needs to keep track of which body is one's own. Only then can the perceptual system assess whether this potentially harmful event is relevant for the subject or not.

Interestingly, it has been found that the brain processes in a specific way what can be conceived of as a margin of safety that immediately surrounds one's body, also known as *peripersonal space*. It was Hediger (1959), the director of the Zurich zoo, who first noted that animals do not process space uniformly and in particular that there is a specific zone immediately surrounding their body that predators cannot approach without eliciting specific responses (flight or fight depending on how close the predator is). As Graziano and Gross (1993, p. 107) described it, peripersonal space is like 'a gelatinous medium surrounding the body that deforms whenever the head rotates or the limbs move'. One of the functions of the representation of peripersonal space is to allocate attentional and motor resources to the area immediately surrounding the body for its protection (for review, see Vignemont and Iannetti, 2015). Now one may ask which body peripersonal space surrounds. Since it is a very primitive mechanism, one may expect that it surrounds the biological body, no matter how one experiences it (as one's own or not). However, we have seen that it is anchored in the body that one takes oneself to have, instead of the body that one actually has, whether it is a rubber hand, a phantom hand, or another person's hand. What matters is that one represents this specific limb as being part of one's own body, and thus worthy of protection. But if one does not represent one's limb as one's own (as in *xenomelia* and *somatoparaphrenia*), then threats lose their personal significance.

To recapitulate, we are aware of our body in a space that goes beyond its boundaries, a space in which there could be threats to which we should immediately react because of their personal significance: they are threats to our body. Within this larger space, it is thus essential to distinguish what is one's own body from what is not. Compare now with pain. The pain that we feel is not space-dependent in the sense that it cannot be more or less close to our body. The space in which we feel pain is the space of the body as we experience it. We do not need to determine whose body is in pain because within the apparent bodily space there is only one body, which is normally our own. Roughly speaking, the enemy is already on the inside. It is thus too late to decide whether it is approaching our body or someone else's body. Paradoxically, pain is too embodied, that is, too focused on one's body, to require a sense of bodily ownership. To some extent, we do not care whose body is hurt; what matters is that we are in pain.

How, then, can we explain pain in the embodiment delusion? In this specific case, indeed it seems that the appropriation of the other person's limb is a prerequisite for the patient to feel pain in it. However, the situation here is abnormal, to say the least. Patients see the other person's left hand aligned with their trunk and do not see their own left hand. This set up leads the other person's hand to become embodied (the patients take it to be their own hand). They then see noxious stimulations applied on the embodied hand. They rate the intensity of the pain as highly as when the noxious stimulations

are applied on their own biological hand. This effect was found only when they appropriated the other person's hand. What is striking here is that the pain is induced *visually*. The fact that vision can hurt may seem surprising, but this is consistent with some results on visuo-tactile interaction in peripersonal space: seeing a visual stimulus close to one's limb can interfere with the perception of a tactile stimulation on the limb (Spence et al., 2004). One way to interpret this cross-modal effect is to say that the sight of objects close to one's body can generate expectation of a tactile event, which then influences the experience of the actual tactile stimulus. Along the same lines, one may suggest that the sight of the noxious stimulus generates expectation of a painful event, which then triggers a painful sensation. What is important here is that the perception of the noxious stimulation is not different from the perception of threats: in both cases, one has visual access to the body perceived in external space. Therefore, one needs to track which body is relevant. The embodiment delusion does not show that the sense of bodily ownership is a necessary condition for normally induced pain. It simply confirms that the external perception of threats – whether they are approaching or already on the skin – requires the body to be singled out as being one's own.

4.2 *Is pain sufficient for the sense of bodily ownership?*

Let us now consider whether pain is sufficient for the awareness of one's body as one's own. I argued in the previous section that pain was too embodied for requiring the sense of ownership, but if it were so embodied, then it should be sufficient to feel pain in a limb to experience this limb as one's own. Yet this is not the case. On the other hand, if one has never felt pain, then one does not experience one's body as one's own. How can one account for this contrastive picture?

Consider the case of phantom pain. One can feel pain to be located in a limb that has been amputated because the missing limb is still mentally represented, and this mental representation is used to spatially organise bodily sensations. One may call it body schema (Head and Holmes, 1911) or long-term body image (O'Shaughnessy, 1980). Here I shall prefer the less theory-laden term of *body map*. Its content carries information about the enduring properties of the body, such as its dimension and its anatomical configuration. The hypothesis is thus that when one feels pain (or any other bodily sensation), one experiences the pain to be located within a frame a reference given by the body map. If the body map includes x , then one can feel sensation in x .

How then can one explain the fact that one can feel pain in a body part that is disowned? It seems that if the body part is disowned, then it is not included in the body map. Yet one feels sensations in it. To solve this apparent puzzle, one solution is to posit more than one type of body map, and thus more than one type of bodily frame. It then becomes possible to have one type of body map that still includes the body part, thus explaining how one can feel sensations there, and another type of body map that no longer includes it, thus explaining that one no longer experiences the body part as one's own. I shall not go here in the details of

the evidence in favour of multiple body representations (see Vignemont, 2017b). Let us simply mention differences in malleability. As said earlier, one can feel tactile sensations in tools. This indicates that the tools have been incorporated in one type of body map, which must be able to adjust very quickly to constant changes. By contrast, the body map that still represents the amputated limb is clearly not easily updated and must thus be less flexible. One might even say that it corresponds to the body by default. I propose that only this latter body map can ground the sense of bodily ownership. I call it the *protective body map* because its function is to keep track of the body necessary for survival (see Vignemont, 2018 for details). To recapitulate, several distinct frames of reference can be used to localise bodily sensations, but only one of them can be at the origin of the sense of bodily ownership. Normally, the location of pain is given within the frame of the protective body map. However, if this representation is disrupted, and if it no longer includes the limb, one can still feel pain in this limb because of the other bodily frames, although one has no sense of ownership for it. Consequently, feeling pain in *x* does not necessarily suffice for one to experience *x* as being part of one's body.

Now the question is how this protective body map is built up. To fix the bodily boundaries that have evolutionary significance, the protective body map must answer two questions: (i) where does the body stop and the rest of the world start? and (ii) what matters for self-preservation? Touch and vision primarily answer the first question, whereas pain primarily answers the second question. Thanks to touch, one can delineate the apparent limits of the body by being aware of the locations at which one can feel tactile sensations and those at which one cannot, as well as by being aware of the resistance of the external world on the skin. Touch is not our only resource for spatial awareness, and one should not neglect the role of vision, which is the main spatial sense, even for the body (Vignemont, 2014). However, the protective body map does not simply represent spatial boundaries – it represents spatial boundaries that have an affective valence. To fix these boundaries involves determining which body has a specific value. It is at this affective level that pain can intervene. It gives a valence to bodily boundaries: it indicates that these boundaries are the ones to care about and to protect if one wants to survive. It thus teaches that it is important to keep track of them. Without pain, without even pain expectation, bodily boundaries are just spatial boundaries. This does not mean that it is only when the body is in danger that we feel our body as our own. I defend a weaker conception, according to which it suffices that it has been in danger at some point to set up a specific valence to the boundaries of the body. Arguably, there is a stage in development in which it is important, possibly even necessary, to experience pain (or any other affectively loaded bodily sensations) in order to give affective significance to the boundaries of the body, a significance that is at the origin of the sense of bodily ownership. In short, past experiences that the boundaries of the body could be at risk make these boundaries the boundaries of the body that we experience as our own now. Patients with congenital insensitivity lack this expertise, which has prevented their protective body map to develop normally.

5 Conclusion

According to the in-my-body conception of pain, pain is inseparable from the sense of bodily ownership. This view is commonly shared but rarely argued for. Here we explored different reasons that could ground this conception, reasons about the nature of bodily ownership and reasons about the spatial and motivational nature of pain. Some reasons are common to all bodily experiences, while others are specific to pain. However, we saw a series of counterexamples that the in-my-body view can hardly account for. In brief, Mrs Gradgrind is not the only one to feel pain somewhere outside what seems to be her own body. Feeling pain in *x* neither requires being aware of *x* as part of one's own body nor suffices to ground the awareness of *x* as part of one's own body.

To be able to localise pain in an 'alien' body part is one thing, but one might question whether this unusual painful sensation is able to motivate protective behaviour. One may, for instance, expect something like pain asymbolia. Unfortunately, there is no detailed investigation of pain behaviour in patients with disownership syndromes. It seems that the patients are not indifferent to their pain (Maravita, 2008; Melzack, 1990): they complain about it, want to stop it, and scream when pinched. The question, however, is whether their sensations motivate body-directed protective behaviours. As seen earlier, one somatoparaphrenic patient asked to have his 'alien' arm removed and put in a cupboard to stop the pain, but to want one's limbs to be dismantled is clearly not a good way to protect one's body. One way to interpret his reaction is to say that he reacts to the unpleasantness of his experience rather than to his physical injury. Pain thus normally plays a dual motivational role, giving reasons for the subject to react both to the sensory and to the affective dimensions of pain. Without the sense of bodily ownership, however, it can play its role only for the affective level. The sense of bodily ownership may thus eventually be important for pain, or at least for its motivational force.

Acknowledgements

I would like to thank David Bain, Michael Brady, and Jennifer Corns for their very helpful comments. Before meeting them, I did not know that working on pain could be so enjoyable. Without them, this paper would not even exist.

Notes

- 1 I may be able to feel *indirectly* the location of your hand if I am holding it while keeping my eyes closed. But then my experience of your hand results from the combination of proprioception and the exteroceptive content of touch.
- 2 At most, pain can grab attention more than any other bodily experiences. One may then say that by attracting attention to the body, pain enhances the saliency or the vividness of the sense of bodily ownership.
- 3 One may reply that the blind man experiences pressure first in his hand, and then only does he project the sensation on the tip of his cane when touching the floor. However, some empirical findings indicate that there is a sense in which he does feel sensation as

being located at the tip of the cane. A well-known effect in psychology is that one has difficulty judging which hand is touched first when the two hands are crossed and they are touched one after the other (Yamamoto and Kitazawa, 2001a). What is interesting is that participants experience the same difficulty if they cross their hands and if they cross two sticks with their hands uncrossed and the two sticks are vibrated one after the other (Yamamoto and Kitazawa, 2001b). This indicates that the vibration is experienced as being located at the tip of the sticks (which were crossed) rather than on the hands that hold them (which were uncrossed).

- 4 One may even suggest that the function of the sense of bodily ownership is to provide reason to accept this specific body as a practical authority (Vignemont, 2018).
- 5 A further problem is that it is not clear how one would interpret the analgesic effects, if they were to be found. Indeed, it is well known that seeing one's own hand in pain alleviates the pain (Longo et al., 2009). Likewise, seeing the rubber hand could modulate the pain if the rubber hand is taken to be one's hand. In this case, the effect would not result from the sense of disownership for the biological hand, but only from the sense of ownership for the rubber hand.
- 6 A similar study was done with patients with somatoparaphrenia, who also displayed no physiological sign of aversive behaviour when their 'alien' limb was threatened (Romano et al., 2014). However, reaction to pain was not analysed in this study.
- 7 Threatening the phantom limb is actually an old measure used to test its reality, which is known as the Abbattucci's choc à blanc: when amputees see that the location at which they feel their phantom limb to be hit, or merely threatened, they show extreme distress, or even pain, and immediately withdraw their phantom limbs, as shown by the movements of their stump (which was never in danger).
- 8 Patients with embodiment delusion also react when they see a threat approaching the other person's hand, but only when they experience it as of their own (Garbarini et al., 2014).

Bibliography

- Ahronheim, J.C. & Gasner, M.R. (1990) 'The Sloganism of Starvation.' *Lancet* 335: 278–279.
- Ailon, N., Charikar, M., & Newman, A. (2008) 'Aggregating Inconsistent Information: Ranking and Clustering.' *Journal of the ACM* 55(5): 1–27.
- Akins, K. (1996) 'Of Sensory Systems and the "Aboutness" of Mental States.' *The Journal of Philosophy* 93(7): 337–372.
- Alfano, M. & Miller, F.G. (forthcoming) 'Dissolving the Placebo Effect: A Case for Pluralism.'
- Allen, C. (2004) Animal Pain. *Nous* 38(4): 617–643.
- Allen, C., Fuchs, P.N., Shriver, A., & Wilson, H.D. (2005) Deciphering Animal Pain. In *Pain: New Essays on the Nature of Pain and the Methodology of Its Study* (pp. 352–366), M. Aydede (ed.). Cambridge, MA: MIT Press.
- Alston, W. (1968) 'Pleasure.' In *The Encyclopedia of Philosophy*, P. Edwards (ed.), New York: Collier-Macmillan.
- Amanzio, M. & Benedetti, F. (1999) 'Neuropharmacological Dissection of Placebo Analgesia: Expectation-Activated Opioid Systems versus Conditioning-Activated Specific Subsystems.' *The Journal of Neuroscience* 19: 484–494.
- Amanzio, M., Benedetti, F., Porro, C.A., Palermo, S., & Cauda, F. (2013) 'Activation Likelihood Estimation Meta-Analysis of Brain Correlates of Placebo Analgesia in Human Experimental Pain.' *Human Brain Mapping* 34: 738–752.
- American Medication Association (1990) 'Persistent Vegetative State and the Decision to Withdraw or Withhold Life Support.' *JAMA* 263: 426–430.
- Anand, K.J.S. (1998) 'Clinical Importance of Pain and Stress in Preterm Neonates.' *Neonatology* 73: 1–9.
- Anchisi, D. & Zanon, M. (2015) 'A Bayesian Perspective on Sensory and Cognitive Integration in Pain Perception and Placebo Analgesia.' *PLoS one* 10: e0117270.
- Anderson, E. (1993) *Value in Ethics and Economics*. Cambridge, MA: Harvard University Press.
- Apkarian, A.V. (2012) 'Chronic Pain and Addiction Pathways.' *Society of Neuroscience Annual Meeting*, New Orleans.
- Armstrong, D. (1962) *Bodily Sensations*. London: Routledge and Paul.
- Armstrong, D.M. (1968) *A Materialist Theory of the Mind*. New York: Humanities Press.
- Arnow, B. A., Desmond, J. E., Banner, L. L., Glover, G. H., Solomon, A., Polan, M. L., Lue, T. F., & Atlas, S.W. (2002) 'Brain Activation and Sexual Arousal in Healthy, Heterosexual Males.' *Brain* 125: 1014–1023.

- Atlas, L.Y., Wielgosz, J., Whittington, R.A., & Wager, T.D. (2013) 'Specifying the Non-Specific Factors Underlying Opioid Analgesia: Expectancy, Attention, and Affect.' *Psychopharmacology* 231(5): 813–23.
- Au Yeung, S.T., Colagiuri, B., Lovibond, P.F., & Colloca, L. (2014) 'Partial Reinforcement, Extinction, and Placebo Analgesia.' *Pain* 155: 1110–1117.
- Auvray, M., Myin, E., & Spence, C. (2010) 'The Sensory-Discriminative and Affective-Motivational Aspects of Pain.' *Neuroscience and Biobehavioral Reviews* 34: 214–223.
- Aydede, M. (2000) 'An Analysis of Pleasure Vis-à-Vis Pain.' *Philosophy and Phenomenological Research* 61(3): 537–570.
- Aydede, M. (2001) 'Naturalism, Introspection, and Direct Realism about Pain.' *Consciousness and Emotion* 2(1): 29–73.
- Aydede, M. (2005) 'The Main Difficulty with Pain.' In *Pain: New Essays on Its Nature and the Methodology of Its Study*, M. Aydede (ed.), Cambridge, MA: MIT Press.
- Aydede, M. (ed.) (2005) *Pain: New Essays on Its Nature and the Methodology of Its Study*. Cambridge, MA: The MIT Press.
- Aydede, M. (2009a) 'Is Feeling Pain the Perception of Something?' *Journal of Philosophy* 106(10): 531–567.
- Aydede, M. (2009b) 'Pain.' In *The Stanford Encyclopedia of Philosophy* (Spring 2013), Edward N. Zalta (ed.). Retrieved from: <http://plato.stanford.edu/archives/spr2013/entries/pain/>.
- Aydede, M. (2014) 'How to Unify Theories of Sensory Pleasure: An Adverbialist Proposal.' *Review of Philosophy and Psychology* 5(1): 119–133.
- Aydede, M. (forthcoming) 'A Contemporary Account of Sensory Pleasure.' In *Pleasure: A History*, Lisa Shapiro (ed.), Oxford: Oxford University Press.
- Aydede, M. & Güzeldere, G. (2005) 'Cognitive Architecture, Concepts, and Introspection: An Information-Theoretic Solution to the Problem of Phenomenal Consciousness.' *Noûs* 39(2): 197–255.
- Aydede, M. & Fulkerson, M. (2013) 'Affective Qualities.' Presentation at the Pacific APA meeting in San Francisco.
- Aydede, M. & Fulkerson, M. (2014) 'Affect: Representationalists' Headache.' *Philosophical Studies* 170: 175–198.
- Baier, A. (2004) 'Feelings that Matter.' In *Thinking about Feeling*, R. Solomon (ed.), Oxford: Oxford University Press.
- Bain, D. (2011) 'The Imperative View of Pain.' *Journal of Consciousness Studies* 18 (9–10): 164–185.
- Bain, D. (2013) 'What makes pains unpleasant?' *Philosophical Studies* 166(1): 69–89.
- Bain, D. (2014) 'Pains that don't hurt.' *Australasian Journal of Philosophy* 92(2): 305–320.
- Baliki, M.N., Geha, P.Y., Fields, H.L., & Apkarian, A.V. (2010) 'Predicting Value of Pain and Analgesia: Nucleus Accumbens Response to Noxious Stimuli Changes in the Presence of Chronic Pain.' *Neuron* 66: 149–160.
- Baliki, M.N., Schnitzer, T.J., Bauer, W.R., Apkarian, V. (2011) 'Brain morphological signatures for chronic pain.' *PLoS ONE* 6: e26010.
- Baliki, M.N., Mansour, A., Baria, A.T., Huang, L., Berger, S.E., Fields, H.L., & Apkarian, A.V. (2013) 'Parceling Human Accumbens into Putative Core and Shell Dissociates Encoding of Values for Reward and Pain.' *The Journal of Neuroscience* 33: 16383–16393.
- Bargh, J.A. and Morsella, E. (2008) 'The Unconscious Mind.' *Perspectives on Psychological Science* 3: 73–79.
- Bargh, J.A., Gollwitzer, P.S., Lee-Chai, A., Barndollar, K., & Trötschel, R. (2001) 'The Automated Will: Unconscious Activation and Pursuit of Behavioral Goals.' *Journal of Personality and Social Psychology* 81: 1004–1027.

- Barrett, L.F. & Bliss-Moreau, E. (2009) 'Affect as a Psychological Primitive.' *Advances in Experimental Social Psychology* 41: 167–218.
- Bartlett, M.S., Littlewort, G.C., Frank, M.G., & Lee, K. (2014) 'Automatic Decoding of Facial Movements Reveals Deceptive Pain Expressions.' *Current Biology* 24: 738–743.
- Barton, J.J. (2003) 'Disorders of Face Perception and Recognition.' *Neurology Clinic* 21: 521–548.
- Bastian, B., Jetten, J., & Fasoli, F. (2011) 'Cleansing the Soul by Hurting the Flesh: The Guilt Reducing Effect of Pain.' *Psychological Science* 22: 334–335.
- Bastian, B., Jetten, J., Hornsey, & M.J. (2014a) 'Gustatory Pleasure and Pain. The Offset of Acute Physical Pain Enhances Responsiveness to Taste.' *Appetite* 72: 150–155.
- Bastian, B., Jetten, J., Hornsey, M.J., & Leknes, S. (2014b) 'The Positive Consequences of Pain: A Biopsychosocial Approach.' *Personality and Social Psychology Review* 18(3): 256–279.
- Bedford, E. (1957) 'Emotions.' *Proceedings of the Aristotelian Society* 57: 281–304.
- Benedetti, F. (2008) 'Mechanisms of Placebo and Placebo-Related Effects Across Diseases and Treatments.' *Annual Review of Pharmacology and Toxicology* 48: 33–60.
- Benedetti, F. (2009) *Placebo Effects: Understanding the Mechanisms in Health and Disease*. New York: Oxford University Press.
- Benedetti, F. (2013) 'Placebo and the New Physiology of the Doctor-Patient Relationship.' *Physiological Reviews* 93: 1207–1246.
- Benedetti, F. (2014) 'Placebo Effects: From the Neurobiological Paradigm to Translational Implications.' *Neuron* 84: 623–637.
- Benedetti, F. & Amanzio M. (2013) 'Mechanisms of the Placebo Response.' *Pulmonary Pharmacology and Therapeutics* 26: 520–523.
- Benedetti, F., Pollo A., Lopiano L., Lanote M., Vighetti S., & Rainero I. (2003) 'Conscious Expectation and Unconscious Conditioning in Analgesic, Motor, and Hormonal Placebo/Nocebo Responses.' *The Journal of Neuroscience* 23(10): 4315–4323.
- Benedetti, F., Mayberg, H.S., Wager, T.D., Stohler, C.S., and Zubieta, J. (2005) 'Neurobiological Mechanisms of the Placebo Effect.' *The Journal of Neuroscience* 25 (45): 10390–10402.
- Benedetti, F., Thoen, W., Blanchard, C., Vighetti, S., & Arduino, C. (2013) 'Pain as a Reward: Changing the Meaning of Pain from Negative to Positive Co-Activates Opioid and Cannabinoid Systems.' *Pain* 154: 361–367.
- Bermúdez, J.L. (2011) 'Bodily Awareness and Self-Consciousness.' In *Oxford Handbook of the Self*, S. Gallagher (ed.), Oxford: Oxford University Press.
- Berna, C., Leknes, S., Holmes, E.A., Edwards, R.R., Goodwin, G.M., & Tracey, I. (2010) 'Induction of Depressed Mood Disrupts Emotion Regulation Neurocircuitry and Enhances Pain Unpleasantness.' *Biological Psychiatry* 67:1083–1090.
- Bernat, J.L. (2006) 'Chronic Disorders of Consciousness.' *Lancet* 367: 1181–1192.
- Bernatzky, G., Presch, M., Anderson, M., & Panksepp, J. (2011) 'Emotional Foundations of Music as a Non-Pharmacological Pain Management Tool in Modern Medicine.' *Neuroscience and Biobehavioral Reviews* 35: 1989–1999.
- Berridge, K. (2004) 'Motivation Concepts in Behavioral Neuroscience.' *Physiology & Behavior* 81(2): 179–209.
- Berridge, K.C. & Valenstein, E.S. (1991) 'What Psychological Process Mediates Feeding Evoked by Electrical Stimulation of the Lateral Hypothalamus?' *Behavioral Neuroscience* 105: 3–14.
- Berridge, K.C. & Robinson, T.E. (1998) 'What is the Role of Dopamine in Reward: Hedonic Impact, Reward Learning, or Incentive Salience?' *Brain Research Reviews* 28: 309–369.

- Berridge, K.C. & Kringelbach, M.L. (2008) 'Affective Neuroscience of Pleasure: Reward in Humans and Animals.' *Psychopharmacology* 199: 457–480.
- Berridge, K.C. & Kringelbach, M.L. (2011) 'Building a Neuroscience of Pleasure and Well-Being.' *Psychology of Well-Being* 1: 1–3.
- Berridge, K.C. & Kringelbach, M.L. (2013) 'Neuroscience of Affect: Brain Mechanisms of Pleasure and Displeasure.' *Current Opinion in Neurobiology* 23: 294–303.
- Berridge, K.C., Flynn, F.W., Schulkin, J., & Grill, H.J. (1984) 'Sodium Depletion Enhances Salt Palatability in Rats.' *Behavioral Neuroscience* 98: 652–660.
- Berthier, M., Starkstein, S., & Leiguarda R. (1988) 'Asymbolia for Pain: A Sensory-Limbic Disconnection Syndrome.' *Annals of Neurology* 24: 41–49.
- Berthier, M., Starkstein, S.E., Nogues, M.A., Robinson, R.G., & Leiguarda, R.C. (1990) 'Bilateral Sensory Seizures in a Patient with Pain Asymbolia.' *Annals of Neurology* 27(1).
- Bingel, U., Wanigasekera, V., Wiech, K., Ni Mhuirchearthaigh, R., Lee, M.C., Ploner, M., & Tracey, I. (2011) 'The Effect of Treatment Expectation on Drug Efficacy: Imaging the Analgesic Benefit of the Opioid Remifentanyl.' *Science Translational Medicine* 3: 70ra14.
- Bishop, F.L., Jacobson, E.E., Shaw, J.R., & Kaptchuk, T.J. (2012) 'Scientific Tools, Fake Treatments, or Triggers for Psychological Healing: How Clinical Trial Participants Conceptualise Placebos.' *Social Science and Medicine* 74(5): 767–774.
- Blackburn, S. (1993) *Essays in Quasi-Realism*. New York: Oxford University Press.
- Blackburn, S. (1998) *Ruling Passions*. New York: Oxford University Press.
- Blair, R.J., Colledge, E., Murray, L., & Mitchell, D.G. (2001) 'A Selective Impairment in the Processing of Sad and Fearful Expressions in Children with Psychopathic Tendencies.' *Journal of Abnormal Child Psychology* 29: 491–498.
- Block, N. (1983) 'Mental Pictures and Cognitive Science.' *Philosophical Review* 92: 499–541.
- Block, N. (1980) 'Troubles with Functionalism.' In *Readings in Philosophy of Psychology Vol. 1* (pp. 268–305), Ned Block (ed.), Cambridge, MA: Harvard University Press.
- Blood, A.J. & Zatorre, R.J. (2001) 'Intensely Pleasurable Responses to Music Correlate with Activity in Brain Regions Implicated in Reward and Emotion.' *Proceedings of the National Academy of Sciences of the United States of America* 98: 11818–11823.
- Boly, M., Faymonville, M.E., Peigneux, P., Lambermont, B., Damas, F., Luxen, A., Lamy, M., Moonen, G., Maquet, P., & Laureys, S. (2005) 'Cerebral Processing of Auditory and Noxious Stimuli in Severely Brain Injured Patients: Differences between VS and MCS.' *Neuropsychological Rehabilitation* 15: 283–289.
- Boly, M., Faymonville, M.E., Schnakers, C., Peigneux, P., Lambermont, B., Phillips, C., Lancellotti, P., Luxen, A., Lamy, M., Moonen, G., Paquet, P., & Laureys, S. (2008) 'Perception of Pain in the Minimally Conscious State with PET Activation: An Observational Study.' *The Lancet Neurology* 7: 1013–1020.
- Boly, M., Garrido, M.I., Gosseries, O., Bruno, M.A., Boveroux, P., Schnakers, C., Massimini, M., Litvak, V., Laureys, S., & Friston, K. (2011) 'Preserved Feedforward but Impaired Topdown Processes in the Vegetative State.' *Science* 332: 858–862.
- Bootzin, R.R. & Bailey, E.T. (2005) 'Understanding Placebo, Nocebo, and Iatrogenic Treatment Effects.' *Journal of Clinical Psychology* 61(7): 871–880.
- Botvinick, M. & Cohen, J. (1998) 'Rubber Hands "Feel" Touch that Eyes See.' *Nature* 391: 756.
- Botvinick, M.M., Nystrom, L.E., Fissell, K., Carter, C.S., & Cohen, J.D. (1999) 'Conflict Monitoring Versus Selection-for-Action in Anterior Cingulate Cortex.' *Nature* 402: 179–181.

- Braam, A.W., Visser, S., Cath, D.C., & Hoogendijk, W.J. (2006) 'Investigation of the Syndrome of Apotemnophilia and Course of a Cognitive-Behavioural Therapy.' *Psychopathology* 39(1): 32–37.
- Brady, M. (2013) 'Pain and the Euthyphro Dilemma.' Paper presented at the Pain Project Conference, University of Glasgow, June 2013.
- Bramble, B. (2011) 'The Distinctive Feeling Theory of Pleasure.' *Philosophical Studies* 162(2): 201–217.
- Brandt, R. (1979) *A Theory of the Good and the Right*. Oxford: Clarendon Press.
- Breiter, H.C., Gollub, R.L., Weisskoff, R.M., Kennedy, D.N., Makris, N., Berke, J.D., Goodman, J.M., Kantor, H.L., Gastfriend, D.R., Riorden, J.P., Mathew, R.T., Rosen, B.R., & Hyman, S.E. (1997) 'Acute Effects of Cocaine on Human Brain Activity and Emotion'. *Neuron* 19: 591–611.
- Brewer, B. (1995) 'Bodily Awareness and the Self.' In *The Body and the Self*, J.L. Bermudez, T. Marcel, & N. Eilan (eds.). Cambridge, MA: MIT Press.
- Brink, D. (1989) *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.
- Broad, C.D. (1930) *Five Types of Ethical Theory*. London: Routledge and Kegan Paul.
- Buchel, C., Geuter, S., Sprenger, C., & Eippert, F. (2014) 'Placebo Analgesia: A Predictive Coding Perspective.' *Neuron* 81: 1223–1239.
- Buhle, J.T., Silvers, J.A., Wager, T.D., Lopez, R., Onyemekwu, C., Kober, H., Weber, J., & Ochsner, K.N. (2014) 'Cognitive Reappraisal of Emotion: A Meta-Analysis of Human Neuroimaging Studies.' *Cereb Cortex* 24(11): 2981–2990.
- Buonanno, F., Harumoto, T., & Ortenzi, C. (2013) 'The Defensive Function of Trichocysts in Paramecium Tetraurelia Against Metazoan Predators Compared with the Chemical Defense of Two Species of Toxin-Containing Ciliates.' *Zoological Science* 30: 255–261.
- Bush, G., Vogt, B.A., Holmes, J., Dale, A.M., Greve, D., Jenike, M.A., & Rosen, B.R. (2002) 'Dorsal Anterior Cingulate Cortex: A Role in Reward-Based Decision Making.' *Proceedings of the National Academy of Sciences* 99: 523–528.
- Byrne, A. (2009) 'Experience and Content.' *The Philosophical Quarterly* 59(236): 429–451.
- Byrne, A. & Tye, M. (2006) 'Qualia Ain't in the Head.' *Nous* 40(2): 241–255.
- Cabanac, M. (1971) 'Physiological Role of Pleasure.' *Science* 173: 1103–1107.
- Cabanac, M. (1979) 'Sensory Pleasure.' *The Quarterly Review of Biology* 54: 1–29.
- Cabanac, M. (1992) 'Pleasure: The Common Currency.' *Journal of Theoretical Biology* 155: 173–200.
- Cacioppo, S., Bianchi-Demicheli, F., Frum C., Pfaus J.G., & Lewis J.W. (2012) 'The Common Neural Multilevel Kernel Density fMRI Analysis.' *Journal of Sexual Medicine* 9: 1048–1054.
- Cannon, W. (1929) *Bodily Changes in Pain, Hunger, Fear and Rage*. New York: Appleton.
- Capelari, E.D., Uribe, C., & Brasil-Neto, J.P. (2009) 'Feeling Pain in the Rubber Hand: Integration of Visual, Proprioceptive, and Painful Stimuli.' *Perception* 38(1): 92.
- Carruthers, P. (1999) 'Sympathy and Subjectivity.' *Australasian Journal of Philosophy* 77: 465–482.
- Carstens, E., Carstens, M.I., Dessirier, J.M., O'Mahony, M., Simons, C.T., Sudo, M., & Sudo, S. (2002) 'It Hurts So Good: Oral Irritation by Spices and Carbonated Drinks and the Underlying Neural Mechanisms.' *Food Qual Prefer* 13:431–443.
- Carver, C.S. & Harmon-Jones, E. (2009) 'Anger is an Approach-Related Affect: Evidence and Implications.' *Psychological Bulletin* 135: 183–204.
- Cassam, Q. (1995) 'Introspection and Bodily Self-Ascription.' In *The Body and the Self*, J.L. Bermudez, T. Marcel, & N. Eilan (eds.). Cambridge, MA: MIT Press.

- Castro, D.C., Chesterman, N.S., Wu, M.K.H., & Berridge, K.C. (2014) 'Two Cortical Hedonic Hotspots: Orbitofrontal and Insular Sites of Sucrose "Liking" Enhancement.' *Ion Neuroscience 2014*, Washington: Society for Neuroscience.
- Cauda, F., Micon, B.M., Sacco, K., Duca, S., D'Agata, F., Geminiani, G., & Canavero, S. (2009) 'Disrupted Intrinsic Functional Connectivity in the Vegetative State.' *Journal of Neurology, Neurosurgery, and Psychiatry* 80: 429–431.
- Chalmers, D. (2006) 'Perception and the Fall from Eden.' In *Perceptual Experience*, T.S. Gendler and J. Hawthorne (eds.), Oxford University Press, 49–125.
- Chapman, W.P., Arrowood, J.G., & Beecher, H.K. (1943) 'The Analgetic Effects of Low Concentrations of Nitrous Oxide Compared in Man with Morphine Sulphate.' *Journal of Clinical Investigation* 22: 871–875.
- Chatelle, C., Majerus, S., Whyte, J., Laureys, S., & Schnakers, C. (2012) 'A Sensitive Scale to Assess Nociceptive Pain in Patients with Disorders of Consciousness.' *Journal of Neurology, Neurosurgery, and Psychiatry* 83: 1233–1237.
- Chelnokova, O., Laeng, B., Eikemo, M., Riegels, J., Loseth, G., Maurud, H., Willoch, F., & Leknes, S. (2014) 'Rewards of Beauty: The Opioid System Mediates Social Motivation in Humans.' *Molecular Psychiatry* 19: 746–747.
- Chiu, P.H., Holmes, A.J., & Pizzagalli, D.A. (2008) 'Dissociable Recruitment of Rostral Anterior Cingulate and Inferior Frontal Cortex in Emotional Response Inhibition.' *Neuroimage* 42: 988–997.
- Clark, A. (2005) 'Painfulness Is Not a Quale.' In *Pain: New Essays on Its Nature and the Methodology of Its Study*, Aydede (ed.), Cambridge, MA: MIT Press.
- Clark, R.E. & Squire, L.R. (1998) 'Classical Conditioning and Brain Systems: The Role of Awareness.' *Science* 280(5360): 77–81.
- Coghill R.C., McHaffie J.G., & Yen, Y. (2003) 'Neural Correlates of Interindividual Differences in the Subjective Experience of Pain.' *Proceedings of the National Academy of Sciences* 100: 8538–8542.
- Cohen, J. & Fulkerson, M. (2014) 'Affect, Rationalization, and Motivation.' *Review of Philosophy and Psychology* 5(1):103–118.
- Coleman, M.R., Rodd, J.M., Davis, M.H., Johnsrude, I.S., Menon, D.K., Pickard, J.D., & Owen, A.M. (2007) 'Do Vegetative Patients Retain Aspects of Language? Evidence from fMRI.' *Brain* 130: 2494–2507.
- Colloca, L. & Finniss, D. (2012) 'Nocebo Effects, Patient-Clinician Communication, and Therapeutic Outcomes.' *JAMA: The Journal of the American Medical Association* 307: 567–568.
- Comer, S.D., Collins, E.D., MacArthur, R.B., Fischman, M.W. (1999) 'Comparison of Intravenous and Intranasal Heroin Self-Administration by Morphine-Maintained Humans.' *Psychopharmacology (Berl)* 143: 327–338.
- Cooper, S.J., Turkish, S. (1989) 'Effects of Naltrexone on Food Preference and Concurrent Behavioral-Responses in Food-Deprived Rats.' *Pharmacological Biochemistry and Behavior* 33: 17–20.
- Corkin, S. & Hebben, N. (1981) 'Subjective Estimates of Chronic Pain before and after Psychosurgery or Treatment in a Pain Unit.' *Pain (Supplement)* 150.
- Corns, J. (2013) 'Unpleasantness, Motivational Oomph, and Painfulness.' *Mind and Language* 29(2): 238–254.
- Cowey, A. (2010) 'The Blindsight Saga.' *Experimental Brain Research* 1: 3–24.
- Craig, A.D. (2003) 'A New View of Pain as a Homeostatic Emotion.' *Trends in Neuroscience* 26: 303–307.
- Craig, A.D. (2009) 'How Do You Feel—Now? The Anterior Insula and Human Awareness.' *Nature Reviews Neuroscience* 10: 59–70.

- Craig, K.D. (1999) 'Emotions and Psychobiology.' In *Textbook of Pain*, 4th Edition, P.D. Wall & R. Melzack (eds.), Edinburgh: Churchill-Livingstone, 331–343.
- Craig, K.D. & Patrick, C.J. (1985) 'Facial Expression During Induced Pain.' *Journal of Personality and Social Psychology* 48: 1080–1091.
- Craig, K.D., Whitfield, M.F., Grunau, R.V., Linton, J., & Hadjistavropoulos, H.D. (1993) 'Pain in the Preterm Neonate: Behavioural and Physiological Indices.' *Pain* 52: 287–299.
- Crisp, R. (2006) *Reasons and the Good*. Oxford: Oxford University Press.
- Crombez, G., Van Ryckeghem, D.M., Eccleston, C., & Van Damme, S. (2013) 'Attentional Bias to Pain-Related Information: A Meta-Analysis.' *Pain* 154: 497–510.
- Cummins, R. (1981) *The Nature of Psychological Explanation*. Cambridge, MA: MIT Press.
- Cutter, B. & Tye, M. (2011) 'Tracking Representationalism and the Painfulness of Pain.' *Philosophical Issues* 21(1): 90–109.
- Cutter, B. & Tye, M. (2014) 'Pains and Reasons: Why It Is Rational to Kill the Messenger.' *The Philosophical Quarterly* 64(256): 423–433.
- D'Arms, J. (2005) 'Two Arguments for Sentimentalism.' *Philosophical Issues* 15.
- D'Arms, J. & Jacobson, D. (2000) 'Sentiment and Value.' *Ethics* 100.
- Damasio, A. (2000) *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York: Houghton Mifflin Harcourt.
- Damasio, A. & Van Hoesen, G.W. (1983) 'Emotional Disturbances Associated with Focal Lesions of the Limbic Frontal Lobe.' In *Neuropsychology of Human Emotion*, K.M. Heilman & P. Satz (eds.), New York: Guilford, 85–110.
- Damasio, A., Damasio H., & Tranel D. (2013) 'Persistence of Feelings and Sentience after Bilateral Damage of the Insula.' *Cerebral Cortex* 23(4): 833–46.
- Danziger, N. & Willer, J.C. (2009) 'Congenital Insensitivity to Pain.' *Revue Neurologique Paris* (1652): 129–136.
- Darwin, C. (1872/1998) *The Expression of the Emotions in Man and Animals*, 3rd Edition. New York: Oxford University Press.
- Davis, K.D., Hutchison, W.D., Lozano, A.M., & Dostrovsky, J.O. (1994) 'Altered Pain and Temperature Perception Following Cingulotomy and Capsulotomy in a Patient with Schizoaffective Disorder.' *Pain* 59(2): 189–199.
- Davis, W.A. (1981) 'Pleasure and Happiness.' *Philosophical Studies* 39(3): 305–317.
- Davis, W.A. (1982) 'A Causal Theory of Enjoyment.' *Mind* 91(362): 240–256.
- Dawel, A., O'Kearney, R., McKone, E., & Palermo, R. (2012) 'Not Just Fear and Sadness: Meta-Analytic Evidence of Pervasive Emotion Recognition Deficits for Facial and Vocal Expressions in Psychopathy.' *Neuroscience & Biobehavioral Reviews* 36: 2288–2304.
- Demertzi, A., Schnakers, C., Ledoux, D., Chatelle, C., Burno, M.-A., Vandaudenhuyse, A., Boly, M., Moonen, G., & Laureys, S. (2009) 'Different Beliefs about Pain Perception in the Vegetative and Minimally Conscious States: A European Survey of Medical and Paramedical Professionals.' *Progress in Brain Research* 177: 329–339.
- Descalzi, G., Li, X.Y., Chen, T., Mercaldo, V., Koga, K., & Zhuo, M. (2012) 'Rapid Synaptic Potentiation within the Anterior Cingulate Cortex Mediates Trace Fear Learning.' *Molecular Brain* 5: 6.
- Descartes, R. (1641/1979) *Les Méditations Métaphysiques*. Paris: Garnier Flammarion.
- Devinsky, O., Morrell, M.J., & Vogt, B.A. (1995) 'Contributions of Anterior Cingulate to Behavior.' *Brain* 118: 279–306.
- de Vignemont, F. (2013) 'The Mark of Bodily Ownership.' *Analysis* 73(4): 643–651.
- de Vignemont, F. (2014) 'A Multimodal Conception of Bodily Awareness.' *Mind* 123(492): 989–1020.

- de Vignemont, F. (2015) 'Pain and Bodily Care: Whose Body Matters?' *Australasian Journal of Philosophy* 93(3): 542–560.
- de Vignemont, F. (2017a) 'Pain and Touch.' *The Monist* 100: 465–477.
- de Vignemont, F. (2017b) *Mind the Body*. Oxford: Oxford University Press.
- de Vignemont, F. (2018) 'Agency and Bodily Ownership: The Bodyguard Hypothesis.' In *The Subject's Matter*, F. de Vignemont & A. Alsmith (eds.), Cambridge, MA: MIT Press.
- de Vignemont, F. & Jacob, P. (2012) 'What It's Like to Feel Another's Pain.' *Philosophy of Science* 79(2): 295–316.
- de Vignemont, F. & Iannetti, G.D. (2015) 'How Many Peripersonal Spaces?' *Neuropsychologia* 70: 327–34.
- Dickens, C. (1854/2003) *Hard Times*. London: Penguin Classics.
- Dokic, J. (2003) 'The Sense of Ownership: An Analogy Between Sensation and Action.' In *Agency and Self-Awareness: Issues in Philosophy and Psychology*, J. Roessler and N. Eilan (eds.), Oxford: Oxford University Press.
- Dong, W.K., Chudler, E.H., Sugiyama, K., Roberts, V.J., & Hayashi, T. (1994) 'Somatosensory, Multisensory, and Task-Related Neurons in Cortical Area 7b (PF) of Unanesthetized Monkeys.' *Journal of Neurophysiology* 72(2): 542–564.
- Dong, W.K., Hayashi, T., Roberts, V.J., Fusco, B.M., & Chudler, E.H. (1996) 'Behavioral Outcome of Posterior Parietal Cortex Injury in the Monkey.' *Pain* 64(3): 579–587.
- Doyle, T.G., Berridge, K.C., & Gosnell, B.A. (1993) 'Morphine Enhances Hedonic Taste Palatability in Rats.' *Pharmacology, Biochemistry, and Behavior* 46: 745–749.
- Dretske, F. (1995) *Naturalizing the Mind*. Cambridge, MA: MIT Press.
- Duncker, K. (1941) 'On Pleasure, Emotion, and Striving.' *Philosophy and Phenomenological Research* 1(4): 391–430.
- Eccles, R. (2006) 'Mechanisms of the Placebo Effect of Sweet Cough Syrups.' *Respiratory Physiology and Neurobiology* 152(3): 340–348.
- Eccleston, C. & Crombez, G. (1999) 'Pain Demands Attention: A Cognitive-Affective Model of the Interruptive Function of Pain.' *Psychological Bulletin* 125: 356–366.
- Egerton, A., Mehta, M.A., Montgomery, A.J., Lappin, J.M., Howes, O.D., Reeves, S.J., Cunningham, V.J., & Grasby, P.M. (2009) 'The Dopaminergic Basis of Human Behaviors: A Review of Molecular Imaging Studies.' *Neuroscience and Biobehavioral Reviews* 33: 1109–1132.
- Ehrsson, H.H. (2009) 'How Many Arms Make a Pair? Perceptual Illusion of Having an Additional Limb.' *Perception* 38: 310–312.
- Ehrsson, H.H., Wiech, K., Weiskopf, N., Dolan, R.J., & Passingham, R.E. (2007) 'Threatening a Rubber Hand That You Feel is Yours Elicits a Cortical Anxiety Response.' *Proceedings of the National Academy of Sciences of the United States of America* 104(23): 9828–9833.
- Eippert, F., Veit, R., Weiskopf, N., Erb, M., Birbaumer, N., & Anders, S. (2007) 'Regulation of Emotional Responses Elicited by Threat-Related Stimuli.' *Human Brain Mapping* 28: 409–423.
- Eippert, F., Bingel, U., Schoell, E.D., Yacubian, J., Klinger, R., Lorenz, J., & Buchel, C. (2009a) 'Activation of the Opioidergic Descending Pain Control System Underlies Placebo Analgesia.' *Neuron* 63: 533–543.
- Eippert, F., Finsterbusch, J., Bingel, U., & Buchel, C. (2009b) 'Direct Evidence for Spinal Cord Involvement in Placebo Analgesia.' *Science* 326: 404.
- Eisenberg, J.B. (2008) 'Schiavo on the Cutting Edge: Functional Brain Imaging and Its Impact on Surrogate End-of-Life Decision-Making.' *Neuroethics* 1: 75–83.

- Eisenberger, N.I. (2011) 'Why Rejection Hurts: What Social Neuroscience Has Revealed about the Brain's Response to Social Rejection.' In *The Oxford Handbook of Social Neuroscience* (pp. 586–598), J. Decety & J. Cacioppo (eds.), New York: Oxford University Press.
- Eisenberger, N.I., Lieberman, M.D., & Williams, K.D. (2003) 'Does Rejection Hurt: An fMRI Study of Social Exclusion.' *Science* 302: 290–292.
- Eisenberger, N.I., Jarcho, J.M., Lieberman, M.D., & Naliboff, B.D. (2006) 'An Experimental Study of Shared Sensitivity to Physical Pain and Social Rejection.' *Pain* 126: 132–138.
- Ekman, P. (1993) 'Facial expression and emotion.' *American Psychologist* 48: 384–392.
- Ekman, P. & Friesen, W. (1978) *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Palo Alto, CA: Consulting Psychologists Press.
- Ekman, P., Davidson, R. J., & Friesen, W. (1990) 'The Duchenne smile: Emotional expression and brain physiology II.' *Journal of Personality and Social Psychology* 58: 342–353.
- Ellingsen, D.M., Wessberg, J., Eikemo, M., Liljencrantz, J., Endestad, T., Olausson, H., & Leknes, S. (2013) 'Placebo Improves Pleasure and Pain through Opposite Modulation of Sensory Processing.' *Proceedings of the National Academy of Sciences of the United States of America* 110: 17993–17998.
- Ellingsen, D.M., Wessberg, J., Chelnokova, O., Olausson, H., Laeng, B., & Leknes, S. (2014) 'In Touch with Your Emotions: Oxytocin and Touch Change Social Impressions while Others' Facial Expressions Can Alter Touch.' *Psychoneuroendocrinology* 39:11–20.
- Ellsberg, D. (1954) 'Classic and Current Notions of Measurable Utility.' *The Economic Journal* 64(255): 528–556.
- Elvemo, N.A., Landro, N.I., Borchgrevink, P.C., & Haberg, A.K. (2015) 'Reward Responsiveness in Patients with Chronic Pain.' *European Journal of Pain* March 2015.
- Etkin, A., Egner, T., & Kalisch, R. (2011) 'Emotional Processing in Anterior Cingulate and Medial Prefrontal Cortex.' *Trends in Cognitive Science* 15: 85–93.
- Farah, M.J. (2008) 'Neuroethics and the Problem of Other Minds: Implications of Neuroscience for the Moral Status of Brain-Damaged Patients and Nonhuman Animals.' *Neuroethics* 1: 9–18.
- Farroni, T., Menon, E., Rigato, S., & Johnson, M.H. (2007) 'The Perception of Facial Expressions in Newborns.' *The European Journal of Developmental Psychology* 4: 2–13.
- Faure, A., Reynolds, S.M., Richard, J.M., & Berridge, K.C. (2008) 'Mesolimbic Dopamine in Desire and Dread: Enabling Motivation to Be Generated by Localized Glutamate Disruptions in Nucleus Accumbens.' *The Journal of Neuroscience* 28: 7184–7192.
- Feldman, F. (1997) 'On the Intrinsic Value of Pleasures.' *Ethics* 107: 448–466.
- Feldman, F. (2004) *Pleasure and the Good Life*. New York: Oxford University Press.
- Feldner, M.T., Hekmat, H., Zvolensky, M.J., Vowles, K.E., Secrist, Z., & Leen-Feldner, E.W. (2006) 'The Role of Experiential Avoidance in Acute Pain Tolerance: A Laboratory Test.' *Journal of Behavior Therapy & Experimental Psychiatry* 37: 146–158.
- Fernandez, E. & Turk, D. (1992) 'Sensory and Affective Components of Pain: Separation and Synthesis.' *Psychological Bulletin* 112: 205–217.
- Fields, H. (2004) 'State-Dependent Opioid Control of Pain.' *Nature Reviews Neuroscience* 5: 565–575.
- Fields, H.L. (1999) 'Pain: An Unpleasant Topic.' *Pain Supplement* 6: S61–S69.
- Fields, H.L. (2006) 'A Motivation-Decision Model of Pain: The Role of Opioids.' In *Proceedings of the 11th World Congress on Pain* pp. (449–459), H. Flor, E. Kalso, & Dostrovsky, J.O. (eds.), Seattle: IASP Press.

- Fields, H.L. (2007) 'Understanding How Opioids Contribute to Reward and Analgesia.' *Regional Anesthesia and Pain Medicine* 32: 242–246.
- Fields, H.L. & Margolis, E.B. (2015) 'Understanding Opioid Reward.' *Trends in Neurosciences* 38(4):217–25.
- Finniss, D.G., Kaptchuk, T.J., Miller, F., & Benedetti, F. (2010) 'Biological, Clinical, and Ethical Advances of Placebo Effects.' *Lancet* 375: 686–695.
- Fins, J.J. (2006) 'Affirming the Right to Care, Preserving the Right to Die: Disorders of Consciousness and Neuroethics after Schiavo.' *Palliative Support Care* 4: 169–178.
- First, M.B. (2005) 'Desire for Amputation of a Limb: Paraphilia, Psychosis, or a New Type of Identity Disorder.' *Psychological Medicine* 35(6): 919–928.
- Fishbain, D.A., Cutler, R., Rosomoff, H.L., & Rosomoff, R.S. (1997) 'Chronic Pain Associated Depression: Antecedent or Consequence of Chronic Pain? A Review.' *Clinical Journal of Pain* 13: 116–137.
- Fletcher, P.C. & Frith, C.D. (2009) 'Perceiving is Believing: A Bayesian Approach to Explaining the Positive Symptoms of Schizophrenia.' *Nature Reviews Neuroscience* 10: 48–58.
- Folegatti, A., de Vignemont, F., Pavani, F., Rossetti, Y., & Farnè, A. (2009) 'Losing One's Hand: Visual-Proprioceptive Conflict Affects Touch Perception.' *PLoS One* 4(9): e6920.
- Foltz, E.L. & White, L.E. (1962) 'Pain "Relief" by Frontal Cingulotomy.' *Journal of Neurosurgery* 19: 89–100.
- Förderreuther, S., Sailer, U. & Straube, A. (2004) 'Impaired Self-Perception of the Hand in Complex Regional Pain Syndrome (CRPS).' *Pain* 110(3): 756–761.
- Frances, A. & Gale, L. (1975) 'The Proprioceptive Body Image in Self-Object Differentiation—A Case of Congenital Indifference to Pain and Head-Banging.' *Psychoanalytic Quarterly* 44(1): 107–126.
- Frank, R.H. (1988) *Passions within Reason: The Strategic Role of the Emotions*. New York: Norton.
- Franklin, K.B.J. (1989) 'Analgesia and the Neural Substrate of Reward.' *Neuroscience and Biobehavioral Reviews* 13: 149–154.
- Freeman, W. & Watts, J.W. (1948) 'Pain Mechanisms and the Frontal Lobes: A Study of Prefrontal Lobotomy for Intractable Pain.' *Annals of Internal Medicine* 28: 747–754.
- Friston, K. & Kiebel, S. (2009) 'Predictive Coding Under the Free-Energy Principle.' *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences* 364: 1211–1221.
- Fullbright, R.K., Troche, C.J., Skudlarski, P., Gore, J.C., & Wexler, B.E. (2001) 'Functional MR Imaging of Regional Brain Activation Associated with the Affective Experience of Pain.' *American Journal of Roentgenology* 177: 1205–1210.
- Fulda, S. & Wetter, T.C. (2008) 'Where Dopamine Meets Opioids: A Meta-Analysis of the Placebo Effect in Restless Legs Syndrome Treatment Studies.' *Brain* 131(4): 902–917.
- Garbarini, F., Fornia, L., Fossataro, C., Pia, L., Gindri, P., & Berti, A. (2014) 'Embodiment of Others' Hands Elicits Arousal Responses Similar to One's Own Hands.' *Current Biology* 24(16): R738–R739.
- Garcia-Larrea, L. (2012) 'The Posterior Insular-Opercular Region and the Search of a Primary Cortex for Pain.' *Clinical Neurophysiology* 42(5):299–313.
- Gazzola, V., Spezio, M.L., Etzel, J.A., Castelli, F., Adolphs, R., & Keysers, C. (2012) 'Primary Somatosensory Cortex Discriminates Affective Significance in Social Touch.' *Proceedings of the National Academy of Sciences of the United States of America* 109: E1657–E1666.

- Georgiadis, J.R., Kortekaas, R., Kuipers, R., Nieuwenburg, A., Pruim, J., Reinders, A.A., & Holstege, G. (2006) 'Regional Cerebral Blood Flow Changes Associated with Clitorally Induced Orgasm in Healthy Women.' *The European Journal of Neuroscience* 24: 3305–3316.
- Geuter, S. & Buchel, C. (2013) 'Facilitation of Pain in the Human Spinal Cord by Nocebo Treatment.' *The Journal of Neuroscience* 33: 13784–13790.
- Giacino, J.T. & Kalmar, K. (1997) 'The Vegetative and Minimally Conscious States: A Comparison of Clinical Features and Functional Outcome.' *Journal of Head Trauma Rehabilitation* 12: 42.
- Giacino, J. T., Ashwal, S., Childs, N., Cranford, R., Jennett, B., Katz, D.I., Kelly, J.P., Rosenberg, J.H., Whyte, J., Zafonte, R.D., & Zasler, N.D. (2002) 'The Minimally Conscious State: Definition and Diagnostic Criteria.' *Neurology* 58: 349–353.
- Gibbard, A. (1990) *Wise Choices, Apt Feelings*. Cambridge, MA: Harvard University Press.
- Gibson, J.J. (1979) *The Ecological Approach to Visual Perception*. Boston, MA: Boston Mifflin.
- Goldstein, I. (1980) 'Why People Prefer Pleasure to Pain.' *Philosophy* 55(213): 349–362.
- González-Franco, M., Peck, T.C., Rodríguez-Fornells, A., & Slater, M. (2013) 'A Threat to a Virtual Hand Elicits Motor Cortex Activation.' *Experimental Brain Research* 232(3): 875–887.
- Gosselin, P., Kirouac, G., & Doré, F.Y. (1995) 'Components and Recognition of Facial Expression in the Communication of Emotion by Actors.' *Journal of Personality and Social Psychology* 68: 83–96.
- Grabenhorst, F., Rolls, E.T., & Bilderbeck, A. (2008) 'How Cognition Modulates Affective Responses to Taste and Flavor: Top-Down Influences on the Orbitofrontal and Pregenual Cingulate Cortices.' *Cerebral Cortex* 18: 1549–1559.
- Grahek, N. (2001) *Feeling Pain and Being in Pain*. Cambridge, MA: MIT Press.
- Grau, J.W. (2002) 'Learning and Memory without a Brain.' In *The Cognitive Animal* (pp. 77–88), M. Bekoff, C. Allen, & G.M. Burghardt (eds.), Cambridge, MA: MIT Press.
- Graziano, M.S.A. & Gross, C.G. (1993) 'A Bimodal Map of Space: Somatosensory Receptive Fields in the Macaque Putamen with Corresponding Visual Receptive Fields.' *Experimental Brain Research* 97: 96–109.
- Graziano, M.S.A. & Kastner, S. (2011) 'Human Consciousness and its Relationship to Social Neuroscience: A Novel Hypothesis.' *Journal of Cognitive Neuroscience* 2: 98–113.
- Green, A.L., Pereira, E.A., & Aziz, T.Z. (2010) 'Deep Brain Stimulation and Pleasure.' In *Pleasures of the Brain* (pp 302–319), Kringelbach, M. L. and Berridge, K. C. (eds.), Oxford: Oxford University Press.
- Greene, C.S., Goddard, G., Macaluso, G.M., & Mauro, G. (2008) 'Topical Review: Placebo Responses and Therapeutic Responses. How Are They Related?' *Journal of Orofacial Pain* 23(2):93–107.
- Greenspan, J.D. & Winfield, J.A. (1992) 'Reversible Pain and Tactile Deficits Associated with a Cerebral Tumor Compressing the Posterior Insula and Parietal Operculum.' *Pain* 50(1):29–39.
- Greenspan, J.D., Lee, R.R., & Lenz, F.A. (1999) 'Pain Sensitivity Alterations as a Function of Lesion Location in the Parasyllvian Cortex.' *Pain* 81(3): 273–282.
- Gregory, N.S., Harris, A.L., Robinson, C.R., Dougherty, P.M., Fuchs, P.N., & Sluka, K.A. (2013) 'An Overview of Animal Models of Pain: Disease Models and Outcome Measures.' *The Journal of Pain* 14(11): 1255–1269.
- Gündel, H., O'Connor, M.F., Littrell, L., Fort, C., & Richard, L. (2003) 'Functional Neuroanatomy of Grief: An fMRI Study.' *Journal of Psychiatry* 160: 1946–1953.

- Gybels, J.M., & Sweet, W.H. (1989) *Neurosurgical Treatment of Persistent Pain*. Basel, Switzerland: Karger.
- Haber, S.N. & Knutson, B. (2010) 'The Reward Circuit: Linking Primate Anatomy and Human Imaging.' *Neuropsychopharmacology* 35: 4–26.
- Haggard, P., Iannetti, G.D., & Longo, M.R. (2013) 'Spatial Sensory Organization and Body Representation in Pain Perception.' *Current Biology* 23(4): R164–R176.
- Hall, W. (1981) 'On "Ratio Scales of Sensory and Affective Verbal Pain Descriptors."' *Pain* 11(1): 101.
- Hall, R. (1989) 'Are Pains Necessarily Unpleasant?' *Philosophy and Phenomenological Research* 49(4): 643–659.
- Hall, R. (2008) 'If It Itches, Scratch!' *Australasian Journal of Philosophy* 86(4):525–535.
- Hamblin, C. (1987) *Imperatives*. Oxford: Basil Blackwell.
- Hardy, J.D., Wolff, H.G., & Goodell, H. (1952) *Pain Sensations and Reactions*. Baltimore, MD: Williams & Wilkins Co.
- Hare, R.M. (1964) 'Pain and Evil.' *Proceedings of the Aristotelian Society*, supp. vol. 38.
- Harkins, S.W., Price, D.D., & Braith, J. (1989) 'Effects of Extraversion and Neuroticism on Experimental Pain, Clinical Pain, and Illness Behavior.' *Pain* 36: 209–218.
- Harman, G. (1990) 'The Intrinsic Quality of Experience.' *Philosophical Perspectives* 4: 31–52.
- Hassoun, N. & Kriegel, U. (2008) 'Consciousness and the Moral Permissibility of Infanticide.' *Journal of Applied Philosophy* 25(1): 44–55.
- Head, H. & Holmes, H.G. (1911) 'Sensory Disturbances from Cerebral Lesions.' *Brain* 34: 102–254.
- Heath, R.G. (1972) 'Pleasure and Brain Activity in Man. Deep and Surface Electroencephalograms during Orgasm.' *The Journal of Nervous and Mental Disease* 154: 3–18.
- Heathwood, C. (2007) 'The Reduction of Sensory Pleasure to Desire.' *Philosophical Studies* 133: 23–44.
- Hediger, H. (1950) *Wild Animals in Captivity*. London: Butterworths Scientific Publications.
- Heinzel, A., Schäfer, R., Müller, H.W., Schieffer, A., Ingenhag, A., Northoff, G., Franz, M., & Hautzel, H. (2010) 'Differential Modulation of Valence and Arousal in High-Alexithymic and Low-Alexithymic Individuals.' *NeuroReport* 21: 998–1002.
- Helm, B. (2001) *Emotional Reason*. Cambridge: Cambridge University Press.
- Helm, B. (2002) 'Felt Evaluations: A Theory of Pleasure and Pain.' *American Philosophical Quarterly* 39(1): 13–30.
- Hemphill, R.E. & Stengel, E. (1940) 'A Study on Pure Word-Deafness.' *Journal of Neurology and Psychiatry* 3: 251–262.
- Hilbert, D. & Klein, C. (2014) 'No Problem.' In *Consciousness Inside and Out* (pp. 299–306), R. Brown (ed.), Oxford: Oxford University Press.
- Hilti, L.M., Hänggi, J., Vitacco, D.A., Kraemer, B., Palla, A., Luechinger, R., Jäncke, L., & Brugger, P. (2013) 'The Desire for Healthy Limb Amputation: Structural Brain Correlates and Clinical Features of Xenomelia.' *Brain* 136(Pt 1): 318–329.
- Ho, C.Y. & Berridge, K.C. (2013) 'An Orexin Hotspot in Ventral Pallidum Amplifies Hedonic 'Liking' for Sweetness.' *Neuropsychopharmacology* 38: 1655–1664.
- Hofbauer, R.K., Rainville, P., Duncan, G.H., & Bushnell, M.C. (2001) 'Cortical Representation of the Sensory Dimension of Pain.' *Journal of Neurophysiology* 86: 402–411.
- Hoffman, G., Harrington, A., & Fields, H. (2005) 'Pain and the Placebo: What We Have Learned.' *Perspectives in Biology and Medicine* 48(2): 248–265.

- Hollin, G.J.S. & Derbyshire, S.W.G. (2009) 'Cold Pressor Pain Reduces Phobic Fear but Fear does not Reduce Pain.' *The Journal of Pain* 10: 1058–1064.
- Horn, C., Blischke, Y., Kunz, M., & Lautenbacher, S. (2012) 'Does Pain Necessarily Have an Affective Component? Negative Evidence from Blink Reflex Experiments.' *Pain Research & Management* 17: 15–24.
- Horne, M. (2009) 'Are People in a Persistent Vegetative State Conscious?' *Monash Bioethics Review* 28(1): 1–12.
- Hosokawa, T., Kato, K., Inoue, M., & Mikami, A. (2007) 'Neurons in the Macaque Orbitofrontal Cortex Code Relative Preference of both Rewarding and Aversive Outcomes.' *Neuroscience Research* 57: 434–445.
- Howard, S. & Hughes, B.M. (2008) 'Expectancies, Not Aroma, Explain Impact of Lavender Aromatherapy on Psychophysiological Indices of Relaxation in Young Healthy Women.' *British Journal of Health Psychology* 13(4): 603–617.
- Hrobjartsson, A. & Gotzsche, P.C. (2001) 'Is the Placebo Powerless? An Analysis of Clinical Trials Comparing Placebo with No Treatment.' *New England Journal of Medicine* 344: 1594–1602.
- Hrobjartsson, A. & Gotzsche, P.C. (2004) 'Is the Placebo Powerless? Update of a Systematic Review with 52 New Randomized Trials Comparing Placebo with No Treatment.' *Journal of Internal Medicine* 256(2): 91–100.
- Hrobjartsson, A. & Gotzsche, P.C. (2010) 'Placebo Interventions for All Clinical Conditions.' *Cochrane Database Systematic Reviews* Jan 20(1): CD003974.
- Hrobjartsson, A., Kaptchuk, T.J., & Miller, F.G. (2011) 'Placebo Effect Studies Are Susceptible to Response Bias and to Other Types of Biases.' *Journal of Clinical Epidemiology* 64: 1223–1229.
- Hughes, S.M. & Nicholson, S.E. (2008) 'Sex Differences in the Assessment of Pain versus Pleasure Facial Expressions.' *Journal of Social, Evolutionary, and Cultural Psychology* 2: 289–298.
- Hull, J.G. & Bond, C.F., Jr. (1986) 'Social and Behavioral Consequences of Alcohol Consumption and Expectancy: A Meta-Analysis.' *Psychological Bulletin* 99: 347–360.
- Hurt, R.W. & Ballantine, H.T.J. (1974) 'Stereotactic Anterior Cingulate Lesions for Persistent Pain: A Report on 68 Cases.' *Clinical Neurosurgery* 21: 334–51.
- Hutchison, W.D., Davis, K.D., Lozano, A.M., Tasker, R.R., & Dostrovsky, J.O. (1999) 'Pain-Related Neurons in the Human Cingulate Cortex.' *Nature Neuroscience* 2(5): 403–405.
- Iannetti, G.D. & Mouraux, A. (2010) 'From the Neuromatrix to the Pain Matrix (and Back).' *Experimental Brain Research* 205: 1–12.
- Iannetti, G.D., Salomons, T.V., Moayed, M., Mouraux, A., & Davis, K.D. (2013) 'Beyond Metaphor: Contrasting Mechanisms of Social and Physical Pain.' *Trends in Cognitive Sciences* 17: 371–378.
- International Association for the Study of Pain (1979) 'Pain Terms: A List with Definitions and Notes on Usage.' *Pain* 6: 249–252.
- International Association for the Study of Pain (2012) 'IASP Taxonomy.' Retrieved 4 May 2014. <https://www.iasp-pain.org/Education/Content.aspx?ItemNumber=1698>.
- International Working Party Report on the Vegetative State (1996) London: Royal Hospital for Neurodisability.
- Jacobson, D. (2011) 'Fitting-Attitude Theories of Value.' *Stanford Encyclopedia of Philosophy*. Retrieved from: <https://plato.stanford.edu/>.
- Jacobson, H. (2013) 'Killing the Messenger: Representationalism and the Painfulness of Pain.' *The Philosophical Quarterly* 63(252): 509–519.

- Jacobsen, P.B., Bovbjerg, D.H., Schwartz, M.D., Andrykowski, M.A., Fetterman, A.D., Gilewski, T., Norton, L., and Redd, W.H. (1993) 'Formation of Food Aversions in Cancer Patients Receiving Repeated Infusions of Chemotherapy.' *Behavior Research and Therapy* 31(8): 739–748.
- James, W. (1884) 'What is an Emotion?' *Mind* 9(34): 188–205.
- Jennett, B. (2002) *The Vegetative State. Medical Facts, Ethical and Legal Dilemmas*. Cambridge: Cambridge University Press.
- Jensen, K.B., Kaptchuk, T., Kirsch, I., Raicek, J., Lindstrom, K.M., Bernab, C., Golluba, R.L., Ingvar, M., & Kong, J. (2012) 'Nonconscious Activation of Placebo and Nocebo Pain Responses.' *Proceedings of the National Academy of Science of the United States of America* 109(39): 15959–15964.
- Jeon, K. (1973) *The Biology of the Amoeba*. New York: Academic Press.
- Johansen, J.P. & Fields, H.L. (2004) 'Glutamatergic Activation of Anterior Cingulate Cortex Produces an Aversive Teaching Signal.' *Nature Neuroscience* 7: 398–403.
- Johnston, M. (2001) 'The Authority of Affect.' *Philosophy and Phenomenological Research* 63(1): 181–214.
- Johnston, C.C., Collinge, J.M., Henderson, S.J., & Anand, K.J. (1997) 'A Cross-Sectional Survey of Pain and Pharmacological Analgesia in Canadian Neonatal Intensive Care Units.' *Clinical Journal of Pain* 13: 308–312.
- Kagan, S. (1992) 'The Limits of Well-Being.' *Social Philosophy and Policy* 9: 169–189.
- Kahane, G. (2009) 'Pain, Dislike and Experience.' *Utilitas* 21(3): 327–336.
- Kahane, G. & Savulescu, J. (2009) 'Brain Damage and the Moral Significance of Consciousness.' *Journal of Medicine and Philosophy* 34: 6–26.
- Kammers, M.P., Rose, K., & Haggard, P. (2011) 'Feeling Numb: Temperature, but Not Thermal Pain, Modulates Feeling of Body Ownership.' *Neuropsychologia* 49(5): 1316–1321.
- Kant, I. (1785) *Grounding for the Metaphysics of Morals*, 3rd Edition. J. W. Ellington (Trs. 1993). New York: Hackett.
- Kappesser, J. & Williams, A.C. (2002) 'Pain and Negative Emotions in the Face: Judgments by Health Care Professionals.' *Pain* 99: 197–206.
- Kaptchuk, T.J. (2011) 'Placebo Studies and Ritual Theory: A Comparative Analysis of Navajo, Acupuncture and Biomedical Healing.' *Philosophical Transactions of the Royal Society of Biological Sciences* 366(1572): 1849–1858.
- Kaptchuk, T.J., Shaw, J., Kerr, C.E., Conboy, L.A., Kelley, J.M., Csordas, T.J., Lembo, A.J., & Jacobson, E.E. (2009) "Maybe I Made Up the Whole Thing": Placebos and Patients' Experiences in a Randomized Controlled Trial.' *Culture, Medicine, and Psychiatry* 33(3): 382–411.
- Kassubek, J., Juengling, F.D., Els, T., Spreer, J., Herpers, M., Krause, T., Moser, E., & Lücking, C.H. (2003) 'Activation of a Residual Cortical Network during Painful Stimulation in Long-Term Postanoxic Vegetative State: A 15O-H₂O PET Study.' *Journal of Neurological Science* 212: 85–91.
- Katz, D. (1925) *The World of Touch*. New York: Psychology Press.
- Keets, A.S. & Beecher, H.K. (1950) 'Pain Relief with Hypnotic Doses of Barbiturates, and a Hypothesis.' *Journal of Pharmacology and Experimental Therapeutics* 100(1): 1–13.
- Kenntner-Mabiala, R. & Pauli, P. (2005) 'Affective Modulation of Brain Potentials to Painful and Nonpainful Stimuli.' *Psychophysiology* 42: 559–567.
- Kersten, D., Mamassian, P., & Yuille, A. (2004) 'Object Perception as Bayesian Inference.' *Annual Review of Psychology* 55: 271–304.
- Kiehl, K.A., Liddle, P.F., & Hopfinger, J.B. (2000) 'Error Processing and the Rostral Anterior Cingulate: An Event-Related fMRI Study.' *Psychophysiology* 37: 216–223.

- Kiehl, K.A., Smith, A.M., Hare, R.D., Mendrek, A., Forster, B.B., & Brink, J. (2001) 'Limbic Abnormalities in Affective Processing by Criminal Psychopaths as Revealed by Functional Magnetic Resonance Imaging.' *Biological Psychiatry* 50: 677–684.
- Kirsch, I. (2005) 'Placebo Psychotherapy: Synonym or Oxymoron?' *Journal of Clinical Psychology* 61: 791–803.
- Klein, C. (2007) 'An Imperative Theory of Pain.' *Journal of Philosophy* 104(10): 517–532.
- Klein, C. (2012) 'Imperatives, Phantom Pains, and Hallucination by Presupposition.' *Philosophical Psychology* 25(6): 917–928.
- Klein, C. (2015a) 'What Pain Asymbolia Really Shows.' *Mind* 124(494): 493–516.
- Klein, C. (2015b) *What the Body Commands: The Imperative Theory of Pain*. Cambridge, MA: MIT Press.
- Knight, L.J., Barbaree, H.E., & Boland, F.J. (1986) 'Alcohol and the Balanced-Placebo Design: The Role of Experimenter Demands in Expectancy.' *Journal of Abnormal Psychology* 95: 335–340.
- Knill, D.C. & Richard, W. (1996) *Perception as Bayesian Inference*. Cambridge: Cambridge University Press.
- Knoll, A.T., Meloni, E.G., Thomas, J.B., Carroll, F.I., & Carlezon, W.A. Jr. (2007) 'Anxiolytic-Like Effects of κ -Opioid Receptor Antagonists in Models of Unlearned and Learned Fear in Rats.' *Journal of Pharmacology and Experimental Therapeutics* 323(3): 838–845.
- Knudsen, L., Petersen, G.L., Nørskov, K.N., Vase, L., Finnerup, N., Jensen, T.S., & Svensson, P. (2011) 'Review of Neuroimaging Studies Related to Pain Modulation.' *Scandinavian Journal of Pain* 2: 108–120.
- Koch, C. (2012) *Consciousness: Confessions of a Romantic Reductionist*. Cambridge, MA: MIT Press.
- Koepp, M.J., Gunn, R.N., Lawrence, A.D., Cunningham, V.J., Dagher, A., Jones, T., Brooks, D.J., Bench, C.J., & Grasby, P.M. (1998) 'Evidence for Striatal Dopamine Release during a Video Game.' *Nature* 393: 266–268.
- Kong, J., Gollub, R.L., Polich, G., Kirsch, I., Laviolette, P., Vangel, M., Rosen, B., & Kaptchuk, T.J. (2008) 'A Functional Magnetic Resonance Imaging Study on the Neural Mechanisms of Hyperalgesic Nocebo Effect.' *The Journal of Neuroscience* 28: 13354–13362.
- Korsgaard, C. (1996) *The Sources of Normativity*. Cambridge: Cambridge University Press.
- Koshi, E.B. & Short, C.A. (2007) 'Placebo Theory and its Implications for Research and Clinical Practice: A Review of the Recent Literature.' *Pain Practice* 7(1): 4–20.
- Koskoff, Y.D. (1949) 'Comment on Dynes and Poppen.' *JAMA*.
- Kringelbach, M.L. (2005) 'The Human Orbitofrontal Cortex: Linking Reward to Hedonic Experience.' *Nature Reviews Neuroscience* 6: 691–702.
- Kringelbach, M.L. (2015) 'The Pleasure of Food: Underlying Brain Mechanisms of Eating and other Pleasures.' *Flavour* 4.
- Kringelbach, M.L. & Berridge, K.C. (2010) 'The Functional Neuroanatomy of Pleasure and Happiness.' *Discovery Medicine* 9: 579–587.
- Kringelbach, M.L. & Berridge, K.C. (2012) 'The Joyful Mind.' *Scientific American* 307: 40–45.
- Kringelbach, M.L. & Phillips, H. (2014) *Emotion. Pleasure and Pain in the Brain*. Oxford: Oxford University Press.
- Kross, E., Berman, M.G., Mischel, W., Smith, E.E., & Wager, T.D. (2011) 'Social Rejection Shares Somatosensory Representations with Physical Pain.' *Proceedings of the National Academy of Sciences USA* 108(15): 6270–6275.

- Kuhn, S. & Gallinat, J. (2012) 'The Neural Correlates of Subjective Pleasantness.' *NeuroImage* 61: 289–294.
- Kulkarni, B., Bentley, D.E., Elliott, R., Youell, P., Watson, A., Derbyshire, S.W., Frackowiak, R.S., Friston, K.J., & Jones, A.K. (2005) 'Attention to Pain Localization and Unpleasantness Discriminates the Functions of the Medial and Lateral Pain Systems.' *The European Journal of Neuroscience* 21: 3133–3142.
- Kunz, M., Prkachin, K., & Lautenbacher, S. (2009) 'The Smile of Pain.' *Pain* 145: 273–275.
- Kunz, M., Lautenbacher, S., Leblanc, N., & Rainville, P. (2012) 'Are Both the Sensory and the Affective Dimensions of Pain Encoded in the Face?' *Pain* 153(2): 350–358.
- Kupers, R., Konings, H., Adriaenen, H., & Gybels, J.M. (1991) 'Morphine Differentially Affects the Sensory and Affective Pain Ratings in Neurogenic and Idiopathic Forms of Pain.' *Pain* 47: 5–12.
- Labukt, I. (2012) 'Hedonic Tone and the Heterogeneity of Pleasure.' *Utilitas* 24(2): 172–199.
- LaGraize, S., Labuda, C., Rutledge, R., Jackson, R. & Fuchs, P. (2004) 'Differential Effect of Anterior Cingulate Cortex Lesion on Mechanical Hypersensitivity and Escape/Avoidance Behavior in an Animal Model of Neuropathic Pain.' *Experimental Neurology* 188: 139–148.
- LaGraize, S., Borzan, J., Peng, Y.B., & Fuchs, P. (2006) 'Selective regulation of pain affect following activation of the opioid anterior cingulate cortex system.' *Experimental Neurology* 197: 22–30.
- Lang, C. (1988) *Ueber Gemuthsbewegungen*, Leipzig: Theodor Thomas.
- Langford, D.J., Crager, S.E., Shehzad, Z., Smith, S.B., Sotocinal, S.G., Levenstadt, J.S., Chanda, M.L., Levitin, D.J., & Mogil, J.S. (2006) 'Social Modulation of Pain as Evidence for Empathy in Mice.' *Science* 312: 1967–1970.
- Larsen, J.T., McGraw, A.P., & Cacioppo, J.T. (2001) 'Can People Feel Happy and Sad at the Same Time?' *Journal of Personality and Social Psychology* 81: 684–696.
- Larsen, J.T., Norris, C.J., & Cacioppo, J.T. (2003) 'Effects of Positive and Negative Affect on Electromyographic Activity over Zygomaticus Major and Corrugator Supercilii.' *Psychophysiology* 40: 776–785.
- Larsen, J.T., McGraw, A.P., Mellers, B.A., & Cacioppo, J.T. (2004) 'The Agony of Victory and Thrill of Defeat: Mixed Emotional Reactions to Disappointing Wins and Relieving Losses.' *Psychological Science* 15: 325–330.
- Larsen, R.J. & Diener, E. (1992) 'Promises and Problems with the Circumplex Model of Emotion.' In *Emotion Review of Personality and Social Psychology*, vol. 13 (pp. 25–59), Clark, M.S. (ed.), Thousand Oaks, CA: Sage Publications, Inc.
- Laureys, S. (2007) 'Eyes Open, Brain Shut.' *Scientific American* 4: 32–37.
- Laureys, S., Lemaire, C., Maquet, P., Phillips, C., & Franck, G. (1999) 'Cerebral Metabolism during Vegetative State and after Recovery to Consciousness.' *Journal of Neurology, Neurosurgery, and Psychiatry* 67: 121.
- Laureys, S., Faymonville, M.E., Peigneux, P., Damas, P., Lambermont, B., Del Fiore, G., Degueldre, C., Aerts, J., Luxen, A., Franck, G., Lamy, M., Moonen, G., & Maquet, P. (2002) 'Cortical Processing of Noxious Somatosensory Stimuli in the Persistent Vegetative State.' *Neuroimage* 17: 732–741.
- Laureys, S., Owen, A.M., & Schiff, N.D. (2004) 'Brain Function in Coma, Vegetative State, and Related Disorders.' *Lancet Neurology* 3: 537–546.
- Lautenbacher, S., Kunderman, B., & Krieg, J. (2006) 'Sleep Deprivation and Pain Perception.' *Sleep Medicine Reviews* 10: 357–369.
- Lazarus, R.S. (1991) *Emotion and Adaptation*. New York: Oxford University Press.

- Lee, S.J., Ralston, H., Drey, E.A., Partridge, J., & Rosen, M.A. (2005) 'Fetal Pain: A Systematic Multidisciplinary Review of the Evidence.' *JAMA* 294(8): 947–954.
- Legrain, V., Iannetti, G.D., Plaghki, L., & Mouraux, A. (2011) 'The Pain Matrix Reloaded: A Salience Detection System for the Body.' *Progress in Neurobiology* 93(1): 111–124.
- Leknes, S. & Tracey, I. (2008) 'A Common Neurobiology for Pain and Pleasure.' *Nature Reviews Neuroscience* 9: 314–320.
- Leknes, S. & Bastian, B. (2014) 'The Benefits of Pain.' *Review of Philosophy and Psychology* 5: 57–70.
- Leknes, S., Berna, C., Lee, M.C., Snyder, G.D., Biele, G., & Tracey, I. (2013) 'The Importance of Context: When Relative Relief Renders Pain Pleasant.' *Pain* 154: 402–410.
- LeResche, L. (1982) 'Facial Expression in Pain: A Study of Candid Photographs.' *Journal of Nonverbal Behavior* 7: 46–56.
- LeResche, L. & Dworkin, S.F. (1988) 'Facial Expressions of Pain and Emotions in Chronic TMD Patients.' *Pain* 35: 71–78.
- Levine, J.D., Gordon, N.C., & Fields, H.L. (1978) 'The Mechanism of Placebo Analgesia.' *Lancet* 2: 654–657.
- Levine, M.E., Stern, R.M., & Koch, K.L. (2006) 'The Effects of Manipulating Expectations through Placebo and Nocebo Administration on Gastric Tachyarrhythmia and Motion-Induced Nausea.' *Psychosomatic Medicine* 68(3): 478–486.
- Lewis M (2000) 'The emergence of human emotions.' *Handbook of Emotions* 2: 265–280.
- Lewis, J.S., Kersten, P., McCabe, C.S., McPherson, K.M., & Blake, D.R. (2007) 'Body Perception Disturbance: A Contribution to Pain in Complex Regional Pain Syndrome (CRPS).' *Pain* 133(1): 111–119.
- Liang, M., Mouraux, A., Hu, L., & Iannetti, G.D. (2013) 'Primary Sensory Cortices Contain Distinguishable Spatial Patterns of Activity for Each Sense.' *Nature Communications* 4: 1979.
- Liccardi, G., Senna, G., Russo, M., Bonadonna, P., Crivellaro, M., Dama, A., D'Amato, M., D'Amato, G., Canonica, G.W., & Passalacqua, G. (2004) 'Evaluation of the Nocebo Effect During Oral Challenge in Patients with Adverse Drug Reactions.' *Journal of Investigational Allergology and Clinical Immunology* 14(2): 104–107.
- Lichstein, L. & Sackett, G.P. (1971) 'Reactions by Differentially Raised Rhesus Monkeys to Noxious Stimulation.' *Developmental Psychobiology* 44: 339–352.
- Liljencrantz, J., Ellingsen, D.M., Leknes, S., Kramer, H., Spadoni, A.D., Kosheleva, E., Simmons, A., Strigo, I., & Olsson, H. (2012) 'C-Tactile Afferent Stimulation Modulate Pain Perception.' In *Society for Neuroscience*, New Orleans, LA.
- Lindquist, K.A., Wager, T.D., Kober, H., Bliss-Moreau, E., & Barrett, L.F. (2012) 'The Brain Basis of Emotion: A Meta-Analytic Review.' *Behavioral and Brain Sciences* 35: 121–143.
- Lombardi, C., Gargioni, S., Canonica, G.W., & Passalacqua, G. (2008) 'The Nocebo Effect During Oral Challenge in Subjects with Adverse Drug Reactions.' *European Annals of Allergy and Clinical Immunology* 40(4): 138–141.
- Longo, M.R., Betti, V., Aglioti, S.M., & Haggard, P. (2009) 'Visually Induced Analgesia: Seeing the Body Reduces Pain.' *The Journal of Neuroscience* 29(39): 12125–12130.
- Lorenz, J., Minoshima, S., & Casey, K.L. (2003) 'Keeping Pain out of Mind: The Role of the Dorsolateral Prefrontal Cortex in Pain Modulation.' *Brain* 126:1079–1091.
- Loseth, G.E., Ellingsen, D.M., & Leknes, S. (2014) 'State-Dependent Mu-Opioid Modulation of Social Motivation.' *Frontiers in Behavioral Neuroscience* 8: 430.
- Lotze, H. (1888) *Microcosmos: An Essay Concerning Man and his Relation to the World*. New York: Scribner & Welford.

- Lu, H.C., Hsieh, J.C., Lu, C.L., Niddam, D.M., Wu, Y.T., Yeh, T.C., Cheng, C.M., Chang, F.Y., & Lee, S.D. (2010) 'Neuronal Correlates in the Modulation of Placebo Analgesia in Experimentally-Induced Esophageal Pain: A 3T-fMRI Study.' *Pain* 148: 75–83.
- Lynn, R. & Eysenck, H.J. (1961) 'Tolerance for Pain, Extraversion and Neuroticism.' *Perceptual and Motor Skills* 12: 161–162.
- Mahler, S.V. & Berridge, K.C. (2012) 'What and When to "Want"? Amygdala-Based Focusing of Incentive Salience Upon Sugar and Sex.' *Psychopharmacology* (Berl) 221: 407–426.
- Mahler, S.V., Smith, K.S., & Berridge, K.C. (2007) 'Endocannabinoid Hedonic Hotspot for Sensory Pleasure: Anandamide in Nucleus Accumbens Shell Enhances "Liking" of a Sweet Reward.' *Neuropsychopharmacology* 32: 2267–2278.
- Mancini, F., Longo, M.R., Kammers, M.P., & Haggard, P. (2011) 'Visual Distortion of Body Size Modulates Pain Perception.' *Psychological Science* 22(3): 325–330.
- Mancini, F., Nash, T., Iannetti, G.D., & Haggard, P. (2014) 'Pain Relief by Touch: A Quantitative Approach.' *Pain* 155: 635–642.
- Maravita, A. (2008) 'Spatial Disorders.' In *Cognitive Neurology: A Clinical Textbook* (pp. 89–118), S.F. Cappa, J. Abutalebi, J.F. Demonet, P.C. Fletcher, P. Garrard (eds.), New York: Oxford University Press.
- Marbach, J.J. & Lund, P. (1981) 'Depression, Anhedonia, and Anxiety in Temporomandibular Joint and Other Facial Pain Syndromes.' *Pain* 11: 73–84.
- Marbach, J.J., Richlin, D.M., & Lipton, J.A. (1983) 'Illness Behavior, Depression and Anhedonia in Myofascial Face and Back Pain Patients.' *Psychotherapy and Psychosomatics* 39: 47–54.
- Marcel, A.J. (1974) 'Perception With and Without Awareness.' Paper presented at the *Experimental Psychology Society Annual Meeting*, Stirling, Scotland.
- Marcel, A.J. (1983) 'Conscious and Unconscious Perception: Experiments on Visual Masking and Word Recognition.' *Cognitive Psychology* 15: 197–237.
- Martin, M.G.F. (1992) 'Sight and Touch.' In *The Content of Experience* (pp. 199–201), T. Crane (Ed.), Cambridge: Cambridge University Press.
- Martin, M.G.F. (1993) 'Sense Modalities and Spatial Properties.' In *Spatial representations*, N. Eilan, R. McCarty and B. Brewer (eds.), Oxford: Oxford University Press.
- Martin, M.G.F. (1995) 'Bodily Awareness: A Sense of Ownership.' In *The Body and the Self*, J.L. Bermudez, T. Marcel, N. Eilan (eds.), Cambridge, MA: MIT Press.
- Martínez, M. (2011) 'Imperative Content and the Painfulness of Pain.' *Phenomenology and the Cognitive Sciences* 10(1): 67–90.
- Martínez, M. (2013) 'The Rationalizing Role of Pains.' Presented at the SSPP meeting in Austin, TX, February 2013.
- Martínez, M. (2015) 'Pains as Reasons.' *Philosophical Studies* 172(9): 2261–2274.
- Martínez, M. & Klein, C. (2016) 'Pain signals are predominantly imperative.' *Biology & Philosophy* 31(2): 283–298.
- Martini, M. (2016) 'Real, Rubber or Virtual: The Vision of "One's Own" Body as a Means for Pain Modulation. A Narrative Review.' *Consciousness and Cognition* 43: 143–51.
- Martucci, K.T., Eisenach, J.C., Tong, C., & Coghill, R.C. (2012) 'Opioid-Independent Mechanisms Supporting Offset Analgesia and Temporal Sharpening of Nociceptive Information.' *Pain* 153: 1232–1243.
- Mauk, M.D., Madden, I.V., Barchas, J.D., & Thompson, R.F. (1982) 'Opiates and Classical Conditioning: Selective Abolition of Conditioned Responses by Activation of Opiate Receptors within the Central Nervous System.' *Proceedings of the National Academy of Sciences of the United States of America* 79: 7598–7602.

- Mayo, L.M., Fraser, D., Childs, E., Momenan, R., Hommer, D.W., de Wit, H., & Heilig, M. (2013) 'Conditioned Preference to a Methamphetamine-Associated Contextual Cue in Humans.' *Neuropsychopharmacology* 38: 921–929.
- Mazzola, L., Isnard, J., Peyron, R., Guenot, M., & Mauguiere, F. (2009) 'Somatotopic Organization of Pain Responses to Direct Electrical Stimulation of the Human Insular Cortex'. *Pain* 146: 99–104.
- Mazzola, L., Isnard, J., Peyron, R., & Mauguiere, F. (2012) 'Stimulation of the Human Cortex and the Experience of Pain: Wilder Penfield's Observations Revisited.' *Brain* 135: 631–640.
- McCullagh, P. (1997) 'Do Fetuses Feel Pain?' *British Journal of Medicine* 314(7076): 302–303.
- McCullagh, P. (2004) *Conscious in a Vegetative State? A Critique of the PVS Concept*. New York: Kluwer Academic Publishers.
- McDowell, J. (1998a) 'Values and Secondary Qualities.' in *Mind, Value, and Reality*, Cambridge, MA: Harvard University Press.
- McDowell, J. (1998b) 'Projection and Truth in Ethics.' In *Mind, Value, and Reality*, Cambridge, MA: Harvard University Press.
- McGlone, F., Olausson, H., Boyle, J.A., Jones-Gotman, M., Dancer, C., Guest, S., & Essick, G. (2012) 'Touching and Feeling: Differences in Pleasant Touch Processing between Glabrous and Hairy Skin in Humans.' *The European Journal of Neuroscience* 35: 1782–1788.
- Meagher, M.W., Arnau, R.C., & Rhudy, J.L. (2001) 'Pain and Emotion: Effects of Affective Picture Modulation.' *Psychosomatic Medicine* 63: 79–90.
- Melzack, R. (1983) 'The McGill Pain Questionnaire.' In *Pain Measurement and Assessment* (pp. 41–70), R. Melzack (ed.), New York: Raven Press.
- Melzack, R. (1987) 'The Short-Form McGill Pain Questionnaire.' *Pain* 30: 191–197.
- Melzack, R. (1989) 'Phantom Limbs, the Self and the Brain.' *Canadian Psychology* 30: 1–16.
- Melzack, R. (1990) 'Phantom Limbs and the Concept of a Neuromatrix.' *Trends in Neuroscience* 13(3): 88–92.
- Melzack, R., & Casey, K.L. (1968) 'Sensory, Motivational and Central Control Determinants of Pain: A New Conceptual Model.' In *The Skin Senses* (pp. 423–443), D. Kenshalo (ed.), Springfield, IL: Thomas.
- Merikle, P.M. & Daneman, M. (1996) 'Memory for Unconsciously Perceived Events: Evidence from Anesthetized Patients.' *Consciousness and Cognition* 5: 525–541.
- Merikle, P.M. & Daneman, M. (1998) 'Psychological Investigations of Unconscious Perception.' *Journal of Consciousness Studies* 5: 5–18.
- Merikle, P.M., Smilek, D., & Eastwood, J.D. (2001) 'Perception Without Awareness: Perspectives from Cognitive Psychology.' *Cognition* 79: 115–134.
- Millan, M.J. (2002) 'Descending Control of Pain.' *Progress in Neurobiology* 66: 355–474.
- Miller, F.G., Colloca, L., & Kaptchuk, T.J. (2009) 'The Placebo Effect: Illness and Interpersonal Healing.' *Perspectives in Biology and Medicine* 52(4): 518–539.
- Miller, J.D., Rausher, S., Hyatt, C.S., Maples, J., & Zeichner, A. (2014) 'Examining the Relations Among Pain Tolerance, Psychopathic Traits, and Violent and Non-Violent Antisocial Behavior.' *Journal of Abnormal Psychology* 123: 205–213.
- Miller, J.M., Vorel, S.R., Tranguch, A.J., Kenny, E.T., Mazzoni, P., van Gorp, W.G., & Kleber, H.D. (2006) 'Anhedonia after a Selective Bilateral Lesion of the Globus Pallidus.' *The American Journal of Psychiatry* 163: 786–788.
- Millikan, R. (1984) *Language, Thought and Other Biological Categories*, Cambridge, MA: MIT Press.

- Mitchell, A., & Boss, B.J. (2002) 'Adverse Effects of Pain on the Nervous Systems of Newborns and Young Children: A Review of the Literature.' *The Journal of Neuroscience Nursing* 34: 228–236.
- Moerman, D. (2002) 'Meaning, Medicine, and the "Placebo Effect."' Cambridge: Cambridge University Press.
- Mohan, R., Jensen, K.B., Petkova, V.I., Dey, A., Barnsley, N., Ingvar, M., & Ehrsson, H.H. (2012) 'No Pain Relief with the Rubber Hand Illusion.' *PloS one* 7(12), e52400.
- Monti, M.M., Vanhaudenhuyse, A., Coleman, M.R., Boly, M., Pickard, J.D., Tshibanda, L., Owen, A.M., & Laureys, S. (2010) 'Willful Modulation of Brain Activity in Disorders of Consciousness.' *New England Journal of Medicine* 362: 579–589.
- Moore, G.E. (1903/1993) *Principia Ethica*, Cambridge: Cambridge University Press.
- Morgan, D., Carter, C.S., DuPree, J.P., Yezierski, R.P., & Vierck, C.J., Jr. (2008) 'Evaluation of Prescription Opioids using Operant-Based Pain Measures in Rats.' *Experimental and Clinical Psychopharmacology* 16(5): 367.
- Moscovitch, M., Winocur, G., & Behrmann, M. (1997) 'What is so Special about Face Recognition? Nineteen Experiments on a Person with Object Agnosia and Dyslexia but Normal Face Perception.' *Cognitive Neuroscience* 9: 555–604.
- Moseley, J.B., O'Malley, K., Petersen, N.J., Menke, T.J., Brody, B.A., Kuykendall, D.H., Hollingsworth, J.C., Ahston, C.M., & Wray, N.P. (2002) 'A Controlled Trial of Arthroscopic Surgery for Osteoarthritis of the Knee.' *New England Journal of Medicine* 347: 81–88.
- Moseley, G.L., Olthof, N., Venema, A., Don, S., Wijers, M., Gallace, A., & Spence, C. (2008) 'Psychologically Induced Cooling of a Specific Body Part Caused by the Illusory Ownership of an Artificial Counterpart.' *PNAS* 105: 13169–13173.
- Moseley, G. L., Gallace, A., & Iannetti, G.D. (2012) 'Spatially Defined Modulation of Skin Temperature and Hand Ownership of Both Hands in patients with Unilateral Complex Regional Pain Syndrome.' *Brain* 135(12): 3676–3686.
- Mounce, C.E., Keogh, E., & Eccleston, C.A. (2010) 'Principal Components Analysis of Negative Affect-Related Constructs Relevant to Pain: Evidence for a Three Component Structure.' *The Journal of Pain* 11: 710–717.
- Mouraux, A., Diukova, A., Lee, M.C., Wise, R.G., & Iannetti, G.D. (2011) 'A Multisensory Investigation of the Functional Significance of the "Pain Matrix."' *NeuroImage* 54: 2237–2249.
- Mulligan, K. (1998) 'From Appropriate Emotions to Values.' *The Monist* 81: 161–188.
- Murray, E.A., O'Doherty, J.P., & Schoenbaum, G. (2007) 'What We Know and Do Not Know About the Functions of the Orbitofrontal Cortex after 20 Years of Cross-Species Studies.' *The Journal of Neuroscience* 27: 8166–8169.
- Nagasako, E.M., Oaklander, A.L., & Dworkin, R.H. (2003) 'Congenital Insensitivity to Pain: An Update.' *Pain* 1013: 213–219.
- Nakashima, S.F., Ukezono, M., Nishida, H., Sudo, R., & Takano, Y. (2015) 'Receiving of Emotional Signal of Pain from Conspecifics in Laboratory Rats.' *Royal Society Open Science* 2.
- Naugle, K.M., Fillingim, R.B., & Riley, J.L. 3rd. (2012) 'A Meta-Analytic Review of the Hypoalgesic Effects of Exercise.' *The Journal of Pain* 13(12): 1139–1150.
- Navratilova, E. & Porreca, F. (2014) 'Reward and Motivation in Pain and Pain Relief.' *Nature Neuroscience* 17: 1304–1312.
- Neubert, M.J., Kincaid, W., & Heinricher, M.M. (2004) 'Nociceptive Facilitating Neurons in the Rostral Ventromedial Medulla.' *Pain* 110: 158–165.

- Newport, R., Pearce, R., & Preston, C. (2010) 'Fake Hands in Action: Embodiment and Control of Supernumerary Limbs.' *Experimental Brain Research* 204: 385–395.
- Nieuwenhuys, R. (2011) 'The Insular Cortex: A Review.' *Progress in Brain Research* 195: 123–163.
- Noordhof, P. (2001) 'In Pain.' *Analysis* 61(2): 95–97.
- Nunn, R. 2009. 'It's Time to Put the Placebo Out of Our Misery.' *British Medical Journal* 338: b1568.
- Nunnally, J.C. (1967) *Psychometric Theory*. New York: McGraw-Hill.
- O'Reilly, J.X., Jbabdi, S., Rushworth, M.F., & Behrens, T.E. (2013) 'Brain Systems for Probabilistic and Dynamic Prediction: Computational Specificity and Integration.' *PLoS Biology* 11:e1001662.
- O'Shaughnessy, B. (1980) *The Will*, Volume 1, Cambridge: Cambridge University Press.
- O'Shaughnessy, B. (2003) *Consciousness and the World*, Oxford: Oxford University Press.
- O'Sullivan, B. & Schroer, R. (2012) 'Painful Reasons: Representationalism as a Theory of Pain.' *The Philosophical Quarterly* 62(249): 737–758.
- Olds, J. (1956) 'Pleasure Centers in the Brain.' *Scientific American* 195: 105–116.
- Olds, J. & Milner, P. (1954) 'Positive Reinforcement Produced by Electrical Stimulation of Septal Area and other Regions of Rat Brain.' *Journal of Comparative and Physiological Psychology* 47: 419–427.
- Ostrowsky, K., Magnin, M., Rylvlin, P., Isnard, J., Guenot, M., & Mauguière, F. (2002) 'Representation of Pain and Somatic Sensation in the Human Insula: A Study of Responses to Direct Electrical Cortical Stimulation.' *Cerebral Cortex* 12(4): 376–385.
- Owen, A.M. (2008) 'Disorders of Consciousness.' *Annals of the New York Academy of Sciences* 1124: 225–238.
- Owen, A.M., Coleman, M.R., Menon, D.K., Johnsrude, I.S., Rodd, J.M., Davis, M.H., Taylor, K., & Pickard, J.D. (2005) 'Residual Auditory Function in Persistent Vegetative State: A Combined PET and fMRI Study.' *Neuropsychological Rehabilitation* 15: 290–306.
- Owen, A.M., Coleman, M.R., Boly, M., Davis, M.H., Laureys, S., and Pickard, J.D. (2006) 'Detecting Awareness in the Vegetative State.' *Science* 313: 1402.
- Pacheco-Lopez, G., Engler, H., Niemi, M.B., & Schedlowski, M. (2006) 'Expectations and Associations that Heal: Immunomodulatory Placebo Effects and its Neurobiology.' *Brain, Behavior, and Immunity* 20(5): 430–436.
- Page, G.G. 2004. 'Are There Long-Term Consequences of Pain in Newborn or Very Young Infants?' *The Journal of Perinatal Education* 13: 10–17.
- Panksepp, J. (1998) *Affective Neuroscience*. New York: Oxford University Press.
- Papaleo, F., Kieffer, B.L., Tabarin, A., & Contarino, A. (2007) 'Decreased Motivation to Eat in μ -Opioid Receptor-Deficient Mice.' *European Journal of Neuroscience* 25: 3398–3405.
- Parfit, D. (2011) *On What Matters, Volume I*. Oxford: Oxford University Press.
- Pautz, A. (2010) 'Do Theories of Consciousness Rest on a Mistake?' *Philosophical Issues* 20(1):333–367.
- Pautz, A. (2014) 'The Real Trouble for Phenomenal Externalists.' In *Consciousness Inside and Out*, Richard Brown (ed.), Oxford University Press: 237–298.
- Peciña, S. & Berridge, K.C. (2005) 'Hedonic Hot Spot in Nucleus Accumbens Shell: Where Do Mu-Opioids Cause Increased Hedonic Impact of Sweetness?' *The Journal of Neuroscience* 25: 11777–11786.
- Peciña, S., Smith, K.S., & Berridge, K.C. (2006) 'Hedonic Hot Spots in the Brain.' *The Neuroscientist* 12(6): 500–511.

- Penfield, W. (1968) 'Engrams in the Human Brain. Mechanisms of Memory.' *Proceedings of the Royal Society of Medicine* 61: 831–840.
- Penfield, W. & Boldrey, E. (1937) 'Somatic Motor and Sensory Representation in the Cerebral Cortex of Man as Studied by Electrical Stimulation.' *Brain* 60: 389–443.
- Perini, I., Bergstrand, S., & Morrison, I. (2013) 'Where Pain Meets Action in the Human Brain.' *The Journal of Neuroscience* 33: 15930–15939.
- Perry, C.J. & Barron, A.B. (2013) 'Neural Mechanisms of Reward in Insects.' *Annual Review of Entomology* 58: 543–562.
- Petkova, V.I. & Ehrsson, H.H. (2008) 'If I Were You: Perceptual Illusion of Body Swapping.' *PLoS One* 3(12): e3832.
- Petrovic, P., Dietrick, T., Fransson, P., Andersson, J., Carlsson, K., & Ingvar, M. (2005) 'Placebo in Emotional Processing-Induced Expectations of Anxiety Relief Activate a General Modulatory Network.' *Neuron* 46(6): 957–969.
- Pettit, P. & Smith, M. (1990) 'Backgrounding Desire.' *Philosophical Review* 99: 565–92.
- Peyron, R., Garcia-Larrea, L., Gregoire, M.C., Costes, N., Convers, P., Lavenne, F., Mauguière, F., Michel, D., & Laurent, B. (1999) 'Haemodynamic Brain Responses to Acute Pain in Humans: Sensory and Attentional Networks.' *Brain* 122: 1765–1780.
- Peyron, R., Laurent, B., & Garcia-Larrea, L. (2000) 'Functional Imaging of Brain Responses to Pain: A Review and Meta-Analysis.' *Neurophysiological Clinic* 20: 263–288.
- Phan, K.L., Wager, T.D., Taylor, S.F., & Liberzon, I. (2002) 'Functional Neuroanatomy of Emotion: A Meta-Analysis of Emotion Activation Studies in PET and fMRI.' *Neuroimage* 16: 331–348.
- Phillips, J.M. & Gatchel, R.J. (2000) 'Extraversion–Introversion and Chronic Pain.' In *Personality Characteristics of Patients with Pain* (pp. 181–202), R.J. Gatchel & J.N. Weisberg (eds.), Washington, DC: American Psychological Association.
- Pia, L., Garbarini, F., Fossataro, C., Forna, L., & Berti, A. (2013) 'Pain and Body Awareness: Evidence from Brain-Damaged Patients with Delusional Body Ownership.' *Frontiers in Human Neuroscience* 7.
- Pitcher, G. (1970) 'Pain Perception.' *Philosophical Review* 79(3): 368–393.
- Ploghaus, A., Narain, C., Beckmann, C.F., Clare, S., Bantick, S., Wise, R., Matthews, P.M., Rawlins, J.N., & Tracey, I. (2001) 'Exacerbation of Pain by Anxiety is Associated with Activity in a Hippocampal Network.' *The Journal of Neuroscience* 21(24): 9896–903.
- Pollatos, O., Herbert, B.M., Füstös, J., Weimer, K., Enck, P., & Zipfel, S. (2011) 'Food Deprivation Sensitizes Pain Perception.' *Journal of Psychophysiology* 26: 1–9.
- Pollo, A., Torre, E., Lopiano, L., Rizzone, M., Lanotte, M., Cavanna, A., Bergamasco, B., & Benedetti, F. (2002) 'Expectation Modulates the Response to Subthalamic Nucleus Stimulation in Parkinsonian Patients.' *NeuroReport* 13: 1383–1386.
- Pollo, A., Carlino E, & Benedetti F. (2008) 'The Top-Down Influence of Ergogenic Placebos on Muscle Work and Fatigue.' *European Journal of Neuroscience* 28(2): 379–388.
- Portenoy, R.K., Jarden, J.O., Sidtis, J.J., Lipton, R.B., Foley, K.M., & Rottenberg, D.A. (1986) 'Compulsive Thalamic Self-Stimulation: A Case with Metabolic, Electrophysiologic and Behavioral Correlates.' *Pain* 27: 277–290.
- Porter, F.L, Wolf, C.M, & Miller, J.P. (1999) 'Procedural Pain in Newborn Infants: The Influence of Intensity and Development.' *Pediatrics* 104: e13.
- Portner, P. (2007) 'Imperatives and Modals.' *Natural Language Semantics* 15(4): 351–383.
- Posner, M.I. (1994) 'Attention: The Mechanisms of Consciousness.' *Proceedings of the National Academy of Science U.S.A.* 91: 7398–7403.
- Price, D.D., McGrath, P.A., Rafii, A., & Buckingham, B. (1983) 'The Validation of Visual Analogue Scales as Ratio Scale Measures for Chronic and Experimental Pain.' *Pain* 17:45–56.

- Price, D.D., Finniss D.G., & Benedetti F. (2008) 'A Comprehensive Review of the Placebo Effect: Recent Advances and Current Thought.' *Annual Review of Psychology* 59: 565–590.
- Prinz, J.J. (2004) *Gut Reactions: A Perceptual Theory of Emotion*. New York: Oxford University Press.
- Prkachin, K.M. (1992) 'The Consistency of Facial Expressions of Pain: A Comparison Across Modalities.' *Pain* 51: 297–306.
- Rabinowicz, W. & Rønnow-Rasmussen, T. (2004) 'The Strike of the Demon: On Fitting Pro-Attitudes and Value.' *Ethics* 114: 391–423.
- Rainville, P., Carrier, B., Hofbauer, R.K., Bushnell, M.C., & Duncan, G.H. (1999) 'Dissociation of Sensory and Affective Dimensions of Pain Using Hypnotic Modulation.' *Pain* 82: 159–171.
- Rainville, P., Duncan, G.H., Price, D.D., Carrier, B., & Bushnell, M.C. (1997) 'Pain Affect Encoded in Human Anterior Cingulate but not Somatosensory Cortex.' *Science* 277: 968–971.
- Ramirez, J.M. & Cabanac, M. (2003) 'Pleasure, the Common Currency of Emotions.' *Annals of the New York Academy of Sciences* 1000: 293–295.
- Rhudy, J.L. & Meagher, M.W. (2000) 'Fear and Anxiety: Divergent Effects on Human Pain Thresholds.' *Pain* 84: 65–75.
- Richard, J.M., Castro, D.C., Difeliceantonio, A.G., Robinson, M.J., & Berridge, K.C. (2013) 'Mapping Brain Circuits of Reward and Motivation: In the Footsteps of Ann Kelley.' *Neuroscience and Biobehavioral Reviews* 37: 1919–1931.
- Rief, W., Bingel, U., Schedlowski, M., Enck, P. (2011) 'Mechanisms Involved in Placebo and Nocebo Responses and Implications for Drug Trials.' *Clinical Pharmacology and Therapeutics* 90: 722–726.
- Robinson, W.S. (2006) 'What Is it Like to Like?' *Philosophical Psychology* 19(6): 743–765.
- Rodriguez-Raecke, R., Niemeier, A., Ihle, K., Ruether, W., & May, A. (2009) 'Brain Gray Matter Decrease in Chronic Pain Is the Consequence and Not the Cause of Pain.' *The Journal of Neuroscience* 29: 13,746–13,750.
- Rodriguez-Raecke, R., Doganci, B., Breimhorst, M., Stankewitz, A., Buchel, C., Birklein, F., & May, A. (2010) 'Insular Cortex Activity is Associated with Effects of Negative Expectation on Nociceptive Long-Term Habituation.' *The Journal of Neuroscience* 30: 11363–11368.
- Rolls, E.T. (2005) *Emotion Explained*. Oxford: Oxford University Press.
- Rolls, E.T., O'Doherty, J., Kringelbach, M.L., Francis, S., Bowtell, R., & McGlone, F. (2003) 'Representations of Pleasant and Painful Touch in the Human Orbitofrontal and Cingulate Cortices.' *Cerebral Cortex* 13: 308–317.
- Romano, D. (2014) *Body Representation Shapes the Responses to Threatening Stimuli*. PhD thesis. University of Milano-Bicocca, Department of Psychology.
- Romano, D., Gandola, M., Bottini, G., & Maravita, A. (2014) 'Arousal Responses to Noxious Stimuli in Somatoparaphrenia and Anosognosia: Clues to Body Awareness.' *Brain* 137(Pt 4): 1213–23.
- Romano, D., Sedda, A., Brugger, P., & Bottini, G. (2015) 'Body Ownership: When Feeling and Knowing Diverge.' *Consciousness and Cognition* 34: 140–148.
- Romer Thomsen, K., Whybrow, P.C., & Kringelbach, M.L. (2015) 'Reconceptualizing Anhedonia: Novel Perspectives on Balancing the Pleasure Networks in the Human Brain.' *Frontiers in Behavioral Neuroscience* 9: 49.
- Rose, J.D. (2002) 'The Neurobehavioral Nature of Fishes and the Question of Awareness and Pain.' *Reviews in Fisheries Science* 10: 1–38.

- Rosenthal, D. & Frank, J.D. (1956) 'Psychotherapy and the Placebo Effect.' *Psychological Bulletin* 53: 294–302.
- Roughan, J.V., Coulter, C.A., Flecknell, P.A., Thomas, H.D., & Sufka, K.J. (2014) 'The Conditioned Place Preference Test for Assessing Welfare Consequences and Potential Refinements in a Mouse Bladder Cancer Model.' *PLoS ONE* 9(8): e103362.
- Rozin, A. & Schiller, D. (1980) 'The Nature and Acquisition of a Preference for Chili Pepper by Humans.' *Motivation and Emotion* 4: 77–101.
- Rozin, P. & Kennel, K. (1983) 'Acquired Preferences for Piquant Foods by Chimpanzees.' *Appetite* 4: 69–77.
- Rozin, P., Gruss, L., & Berk, G. (1979) 'Reversal of Innate Aversions: Attempts to Induce a Preference for Chili Peppers in Rats.' *Journal of Comparative and Physiological Psychology* 93: 1001–1014.
- Rozin, P., Guillot, L., Fincher, K., Rozin, A., & Tsukayama, E. (2013) 'Glad to Be Sad, and Other Examples of Benign Masochism.' *Judgment and Decision Making* 8: 439–447.
- Ryle, G. (1954) 'Pleasure.' *Proceedings of the Aristotelian Society*, supp. vol. 28: 135–146.
- Sacks, O. (1984) *A Leg to Stand On*. London: Picador.
- Salamone, J.D. & Correa, M. (2012) 'The Mysterious Motivational Functions of Mesolimbic Dopamine.' *Neuron* 76: 470–485.
- Salimpoor, V.N., Benovoy, M., Larcher, K., Dagher, A., & Zatorre, R.J. (2011) 'Anatomically Distinct Dopamine Release During Anticipation and Experience of Peak Emotion to Music.' *Nature Neuroscience* 14:257–262.
- Sawamoto, N., Honda, M., Okada, T., Hanakawa, T., Kanda, M., Fukuyama, H., Konishi, J., & Shibasaki, H. (2000) 'Expectation of Pain Enhances Responses to Nonpainful Somatosensory Stimulation in the Anterior Cingulate Cortex and Parietal Operculum/Posterior Insula: An Event-Related Functional Magnetic Resonance Imaging Study.' *The Journal of Neuroscience* 20(19): 7438–7448.
- Schilder, P. & Stengel, E. (1928) 'Schmerzasympolie. Zeitschrift fur die gesamte.' *Neurologie und Psychiatri* 113(1):143–158.
- Schmack, K., Gomez-Carrillo de Castro, A., Rothkirch, M., Sekutowicz, M., Rossler, H., Haynes, J.D., Heinz, A., Petrovic, P., & Sterzer, P. (2013) 'Delusions and the Role of Beliefs in Perceptual Inference.' *The Journal of Neuroscience* 33: 13701–13712.
- Schmidt, L., Braun, E.K., Wager, T.D., & Shohamy, D. (2014) 'Mind Matters: Placebo Enhances Reward Learning in Parkinson's Disease.' *Nature Neuroscience*.
- Schnakers, C., and Zasler, N.D. (2007) 'Pain Assessment and Management in Disorders of Consciousness.' *Current Opinion in Neurology* 20: 620–626.
- Schnakers, C., Faymonville, M.-E., & Laureys, S. (2007) 'Ethical Implications: Pain, Coma, and Related Disorders.' In *Encyclopedia of Consciousness*. Retrieved from http://www.credoreference.com/entry/estcon/ethical_implications_pain_coma_and_related_disorders.
- Schnakers, C., Vanhaudenhuyse, A., Giacino, J., Ventura, M., Boly, M., Majerus, S., Moonen, G., & Laureys, S. (2009) 'Diagnostic Accuracy of the Vegetative and Minimally Conscious State: Clinical Consensus versus Standardized Neurobehavioral Assessment.' *BMC Neurology* 9: 35.
- Schnakers, C., Chatelle, C., Vanhaudenhuyse, A., Majerus, S., Ledoux, D., Boly, M., Bruno, M.-A., Boverous, P., Demertzi, A., Moonen, G., & Laureys, S. (2010) 'The Nociception Coma Scale: A New Tool to Assess Nociception in Disorders of Consciousness.' *Pain* 148: 215–219.
- Schnakers, C., Chatelle, C., Demertzi, A., Majerus, S., & Laureys, S. (2012) 'What About Pain in Disorders of Consciousness?' *The AAPS Journal* 14: 437–444.

- Schnitzler, A. & Ploner, M. (2000) 'Neurophysiology and Functional Neuroanatomy of Pain Perception.' *Journal of Clinical Neurophysiology* 17: 592–603.
- Schroeder, T. (2004) *Three Faces of Desire*, New York: Oxford University Press.
- Schulz, K.P., Clerkin, S.M., Halperin, J.M., Newcorn, J.H., Tang, C.Y., & Fan, J. (2009) 'Dissociable Neural Effects of Stimulus Valence and Preceding Context during the Inhibition of Responses to Emotional Faces.' *Human Brain Mapping* 30: 2821–2833.
- Schultz W (2002) 'Getting Formal with Dopamine and Reward.' *Neuron* 36: 241–263.
- Schwartz, N., Temkin, P., Jurado, S., Lim, B.K., Heifets, B.D., Polepalli, J.S., & Malenka, R.C. (2014) 'Chronic Pain. Decreased Motivation during Chronic Pain Requires Long-Term Depression in the Nucleus Accumbens.' *Science* 345: 535–542.
- Scott, D.J., Stohler, C.S., Egnatuk, C.M., Wang, H., Koeppe, R.A., & Zubieta, J.K. (2007) 'Individual Differences in Reward Responding Explain Placebo-Induced Expectations and Effects.' *Neuron* 55: 325–336.
- Scott, D.J., Stohler, C.S., Egnatuk, C.M., Wang, H., Koeppe, R.A., & Zubieta, J.K. (2008) 'Placebo and Nocebo Effects Are Defined by Opposite Opioid and Dopaminergic Responses.' *JAMA Psychiatry* 65: 220–231.
- Scott, W., Trost, Z., Bernier, E., & Sullivan, M.J.L. (2013) 'Anger Differentially Mediates the Relationship between Perceived Injustice and Chronic Pain Outcomes.' *Pain* 154(9): 1691–1698.
- Segerdahl, A.R., Mezue, M., Okell, T.W., Farrar, J.T., & Tracey, I. (2015) 'The Dorsal Posterior Insula Suberves a Fundamental Role in Human Pain.' *Nature Neuroscience*.
- Seminowicz, D.A., Wideman, T.H., Naso, L., Hatami-Khoroushahi, Z., Fallatah, S., Ware, M.A., Jarzem, P., Bushnell, M.C., Shir, Y., Ouellet, J.A., & Stone, L.S. (2011) 'Effective Treatment of Chronic Low Back Pain in Humans Reverses Abnormal Brain Anatomy and Function.' *The Journal of Neuroscience* 31: 7540–7550.
- Seymour, B., Daw, N., Dayan, P., Singer, T., & Dolan, R. (2007) 'Differential Encoding of Losses and Gains in the Human Striatum.' *Journal of Neuroscience* 27: 4826–4831.
- Shackman, A.J., Salomons, T.V., Slagter, H.A., Fox, A.S., Winter, J.J., & Davidson, R.J. (2011) 'The Integration of Negative Affect, Pain and Cognitive Control in the Cingulate Cortex.' *Nature Reviews Neuroscience* 12(3): 154–167.
- Shewmon, D.A., Holmes, G.L., Byrne, P.A. (1999) 'Consciousness in Congenitally Decorticate Children: Developmental Vegetative State as Self-Fulfilling Prophecy.' *Developmental medicine and child neurology* 41: 364–374.
- Shizgal, P., Fulton, S., & Woodside, B. (2001) 'Brain Reward Circuitry and the Regulation of Energy Balance.' *International Journal of Obesity and Related Metabolic Disorders* 25(Suppl 5): S17–S21.
- Shoemaker, S. (1968) 'Self-Reference and Self-Awareness.' *The Journal of Philosophy* 65: 555–567.
- Shriver, A. (2006) 'Minding Mammals.' *Philosophical Psychology* 19: 433–442.
- Shriver, A. (2014) 'What We Can, And Cannot, Learn from a Neural Signature of Pain (abstract).' *American Journal of Bioethics-Neuroscience* 5(3):48.
- Sidgwick, H. (1907/1981) *The Methods of Ethics*, 7th Ed, Macmillan.
- Sidis, B. (1989) *The Psychology of Suggestion*, New York: D. Appleton.
- Simon, D., Craig, K.D., Gosselin, F., Belin, P., & Rainville, P. (2008) 'Recognition and Discrimination of Prototypical Dynamic Expressions of Pain and Emotions.' *Pain* 135: 55–64.
- Singer, P. (1975) *Animal Liberation*. New York: Harper Perennial Modern Classics.
- Skvortsova, V., Palminteri, S., & Pessiglione, M. (2014) 'Learning To Minimize Efforts versus Maximizing Rewards: Computational Principles and Neural Correlates.' *The Journal of Neuroscience* 34: 15621–15630.

- Small, D.M., Zatorre, R.J., Dagher, A., Evans, A.C., & Jones-Gotman, M. (2001) 'Changes in Brain Activity Related to Eating Chocolate: From Pleasure to Aversion.' *Brain: a Journal of Neurology* 124: 1720–1733.
- Smith, K.S. (2010) 'Hedonic Hotspots: Generating Sensory Pleasure in the Brain'. In *Pleasures of the Brain* (pp. 27–49), Kringelbach, M.L. and Berridge, K.C. (eds.), Oxford: Oxford University Press.
- Smith, K.S. & Berridge, K.C. (2007) 'Opioid Limbic Circuit for Reward: Interaction between Hedonic Hotspots of Nucleus Accumbens and Ventral Pallidum.' *The Journal of Neuroscience* 27: 1594–1605.
- Smith, K.S., Berridge, K.C., & Aldridge, J.W. (2011) 'Disentangling Pleasure from Incentive Salience and Learning Signals in Brain Reward Circuitry.' *Proceedings of the National Academy of Sciences of the United States of America* 108: E255–E264.
- Smuts, A. (2011) 'The Feels Good Theory of Pleasure. *Philosophical Studies* 155(2): 241–265.
- Snyder, C.R., Berg, C., Woodward, J.T., Gum, A., Rand, K.L., Wroblewski, K.K., Brown, J., & Hackman, A. (2005) 'Hope against the Cold: Individual Differences in Trait Hope and Acute Pain Tolerance on the Cold Pressor Task.' *Journal of Personality* 73(2): 287–312.
- Soderpalm, A.H., Berridge, K.C. (2000) 'The Hedonic Impact and Intake of Food Are Increased by Midazolam Microinjection in the Parabrachial Nucleus.' *Brain Research* 877: 288–297.
- Spence, C., Pavani, F., & Driver, J. (2004) 'Spatial Constraints on Visual-Tactile Cross-Modal Distractor Congruency Effects.' *Cognitive, Affective, and Behavioral Neuroscience* 4(2): 148–169.
- Spinoza, B. (1677/1992) *Ethics* (S. Shirley, trans.). Indianapolis, IN: Hackett.
- Sprenger, C., Eippert, F., Finsterbusch, J., Bingel, U., Rose, M., & Buchel, C. (2012) 'Attention Modulates Spinal Cord Responses to Pain.' *Current Biology* 22: 1019–1022.
- Strigge, T.L.S. (2000) 'Is the Esse of Intrinsic Value Percipi?: Pleasure, Pain and Value. *Royal Institute of Philosophy* (Supp. Volume) 47: 119–140.
- Spunt, R.P., Lieberman, M.D., Cohen, J.R., & Eisenberger, N.I. (2012) 'The Phenomenology of Error Processing: The Dorsal ACC Response to Stop-Signal Errors Tracks Reports of Negative Affect.' *Journal of Cognitive Neuroscience* 24(8): 1753–1765.
- Stalnaker, R. (2002) 'Common Ground.' *Linguistics and Philosophy* 25: 701–721.
- Starr, C.J., Sawaki, L., Wittenberg, G.F., Burdette, J.H., Oshiro, Y., Quevedo, A.S., & Coghill, R.C. (2009) 'Roles of the Insular Cortex in the Modulation of Pain: Insights from Brain Lesions.' *The Journal of Neuroscience* 29: 2684–2694.
- Steiner, J.E. (1973) 'The Gustofacial Response: Observation on Normal and Anencephalic Newborn Infants.' *Symposium on Oral Sensation and Perception* 254–278.
- Steiner, J.E., Glaser, D., Hawilo, M.E., & Berridge, K.C. (2001) 'Comparative Expression of Hedonic Impact: Affective Reactions to Taste by Human Infants and Other Primates.' *Neuroscience and Biobehavioral Reviews* 25:53–74.
- Stephens, R. & Allsop, C. (2012) 'Effect of Manipulated State Aggression on Pain Tolerance.' *Psychological Reports: Disability & Trauma* 111: 311–321.
- Stephens, R., Atkins, J., & Kingston, A. (2009) 'Swearing as a Response to Pain.' *NeuroReport* 20: 1056–1060.
- Stevens, B., McGrath, P., Gibbins, S., Beyene, J., Breau, L., Camfield, C., Finley, A., Franch, L. Howlett, A., KcKeever, P., O'Brian, K., Ohisson, A., & Yamada, J. (2003) 'Procedural Pain in Newborns at Risk for Neurologic Impairment.' *Pain* 105: 27–35.
- Stewart-Williams, S. & Podd, J. (2004) 'The Placebo Effect: Dissolving the Expectancy Versus Conditioning Debate.' *Psychological Bulletin* 130(2): 324–340.

- Stockhorst, U., Steingrueber, H., Enck, P., & Klosterhalfen, S. (2006) 'Pavlovian Conditioning of Nausea and Vomiting.' *Autonomic Neuroscience: Basic and Clinical* 129: 50–57.
- Sufka, K.J. (1994) 'Conditioned Place Preference Paradigm: A Novel Approach for Analgesic Drug Assessment Against Chronic Pain.' *Pain* 58: 355–366.
- Taddio, A., Katz, J., Ilersich, A.L., & Koren, G. (1997) 'Effect of Neonatal Circumcision on Pain Response During Subsequent Routine Vaccination.' *Lancet* 349: 599–603.
- Taylor, J.J., Borckardt, J.J., & George, M.S. (2012) 'Endogenous Opioids Mediate Left Dorsolateral Prefrontal Cortex rTMS-Induced Analgesia.' *Pain* 153:1219–1225.
- Tindell, A.J., Smith, K.S., Pecina, S., Berridge, K.C., & Aldridge, J.W. (2006) 'Ventral Pallidum Firing Codes Hedonic Reward: When a Bad Taste Turns Good.' *Journal of Neurophysiology* 96: 2399–2409.
- Tölle, T.R., Kaufmann, T., Siessmeier, T., Lautenbacher, S., Berthele, A., Munz, F., Ziegler, W., Willoch, F., Schwaiger, M., Conrad, B., & Bartenstein, P. (1999) 'Region-Specific Encoding of Sensory and Affective Components of Pain in the Human Brain: A Positron Emission Tomography Correlation Analysis.' *Annals of Neurology* 45: 40–47.
- Tracey, I. (2010) 'Getting the Pain you Expect: Mechanisms of Placebo, Nocebo and Reappraisal Effects in Humans.' *Nature Medicine* 16: 1277–1283.
- Tracey, I. & Mantyh, P.W. (2007) 'The Cerebral Signature for Pain Perception and its Modulation.' *Neuron* 55: 377–391.
- Tsakiris, M. (2010) 'My Body in the Brain: A Neurocognitive Model of Body-Ownership.' *Neuropsychologia* 48: 703–712.
- Turner, A. (2012) 'Placebos and the Logic of Placebo Comparison.' *Biology and Philosophy* 27(3): 419–432
- Tye, M. (1995) *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*. Cambridge, MA: MIT Press.
- Tye, M. (2002) 'On the Location of a Pain.' *Analysis* 62(2):150–153.
- Tye, M. (2005) 'Another Look at Representationalism about Pain.' In Aydede (ed.) (2005).
- Urban, M.O., Gebhart, G.F. (1999) 'Supraspinal Contributions to Hyperalgesia.' *Proceedings of the National Academy of Sciences of the United States of America* 96: 7687–7692.
- Valenstein, E.S. (1986) *Great and Desperate Cures: The Rise and Decline of Psychosurgery and other Radical Treatments for Mental Illness*. New York: Basic Books.
- Valenstein, E.S., Cox, V.C., & Kakolewski, J.W. (1970) 'Reexamination of the Role of the Hypothalamus in Motivation.' *Psychological Review* 77: 16–31.
- Valenzuela-Moguillansky, C., O'Regan, J.K., & Petitmengin, C. (2013) 'Exploring the Subjective Experience of the "Rubber Hand" Illusion.' *Frontiers in Human Neuroscience* 7.
- Vallar, G. & Ronchi, R. (2009) 'Somatoparaphrenia: A Body Delusion. A Review of the Neuropsychological Literature.' *Experimental Brain Research* 192(3): 533–551.
- van Laarhoven, A.I.M., Vogelaar, M.L., Wilder-Smith, O.H., van Riel, P.L.C.M., van de Kerkhof, P.C.M., Kraaijaat, F.W., & Evers, A.W.M. (2011) 'Induction of Nocebo and Placebo Effects on Itch and Pain by Verbal Suggestions.' *Pain* 152(7): 1486–1494.
- van Middendorp, H., Lumley, M.A., Jacobs, J.W.G., Bijlsma, J.W.J., & Geenen, R. (2010) 'The Effects of Anger and Sadness on Clinical Pain Reports and Experimentally-Induced Pain Thresholds in Women With and Without Fibromyalgia.' *Arthritis Care & Research* 62L: 1370–1376.

- Vance, J. (2014) 'Emotion and the New Epistemic Challenge from Cognitive Penetrability.' *Philosophical Studies* 169: 257–83.
- Vancleef, L.M.G., Vlaeyen, J.W.S., & Peters, M. (2009) 'Dimensional and Componential Structure of a Hierarchical Organisation of Pain-Related Anxiety Constructs.' *Psychological Assessment* 23: 340–351.
- Vanhoudenhuysse, A., Noirhomme, Q., Tshibanda, L.J., Bruno, M.A., Boveroux, P., Schnakers, C., Soddu, A., Perlberg, V., Ledoux, D., Brichant, J.F., Moonen, G., Greicius, M.D., Laureys, C., & Boly, M. (2010) 'Default Network Connectivity Reflects the Level of Consciousness in Non-Communicative Brain Damaged Patients.' *Brain* 133: 161–171.
- Veehof, M.M., Oskam, M.J., Schreurs, K.M.G., & Bohlmeijer, E.T. (2011) 'Acceptance-Based Interventions for the Treatment of Chronic Pain. A Systematic Review and Meta-analysis.' *Pain* 15: 533–542.
- Veldhuijzen, D.S., Greenspan, J.D., Kim, J.H., & Lenz, F.A. (2010) 'Altered Pain and Thermal Sensation in Subjects with Isolated Parietal and Insular Cortical Lesions.' *European Journal of Pain* 14(5): 535.e1–535.11.
- Vesey, G.N.A. (1961) 'The Location of Bodily Sensations.' *Mind* 277: 25–35.
- Vierk, C.J., Hansson, P.T., & Yeziarski, R.P. (2008) 'Clinical and Pre-Clinical Pain Assessment: Are We Measuring the Same Thing?' *Pain* 135: 7–10.
- Vogt, B. (2005) 'Pain and Emotion Interactions in Subregions of the Cingulate Gyrus.' *Nature Reviews Neuroscience* 6(7): 533–44.
- Volders, S., Meulders, A., De Peuter, S., & Vlaeyen, J.W. (2014) 'The Reduction of Fear of Movement-Related Pain: Does Motivational Context Matter?' *The Clinical Journal of Pain*.
- Volkow, N.D., Wang, G.J., Fowler, J.S., Logan, J., Gatley, S.J., Hitzemann, R., Chen, A.D., Dewey, S.L., & Pappas, N. (1997) 'Decreased Striatal Dopaminergic Responsiveness in Detoxified Cocaine-Dependent Subjects.' *Nature* 386: 830–833.
- Vollm, B.A., de Araujo, I.E., Cowen, P.J., Rolls, E.T., Kringelbach, M.L., Smith, K.A., Jezzard, P., Heal, R.J., & Matthews, P.M. (2004) 'Methamphetamine Activates Reward Circuitry in Drug Naive Human Subjects.' *Neuropsychopharmacology* 29: 1715–1722.
- Vranas, P. (2008) 'New Foundations for Imperative Logic I: Logical Connectives, Consistency, and Quantifiers.' *Noûs* 42(4):529–572.
- Vuust, P. & Kringelbach, M.L. (2010) 'The Pleasure of Making Sense of Music.' *Interdisciplinary Science Reviews* 35: 168–185.
- Waber, R., Shiv, B., Carmon, Z., & Ariely, D. (2008) 'Commercial Features of Placebo and Therapeutic Efficacy.' *Journal of American Medical Association* 299(9): 1016–1017.
- Wager, T.D., Rilling, J.K., Smith, E.E., Sokolik, A., Casey, K.L., Davidson, R.J., Kosslyn, S.M., Rose, R.M., & Cohen, J.D. (2004) 'Placebo-Induced Changes in fMRI in the Anticipation and Experience of Pain.' *Science* 303: 1162–1167.
- Wager, T.D., Atlas, L.Y., Lindquist, M.A., Roy, M., Woo, C., & Kross, E. (2013) 'An fMRI-Based Neurologic Signature of Physical Pain.' *New England Journal of Medicine* 368: 1388–1397.
- Wall, P.D. (1967) 'The Laminar Organization of Dorsal Horn and Effects of Descending Impulses.' *The Journal of Physiology* 188: 403–423.
- Watson, D. & Tellegen, A. (1985) 'Toward a Consensual Structure of Mood.' *Psychological Bulletin* 98: 219–235.
- Wei, F. & Zhuo, M. (2006) 'Potentiation of Sensory Responses in the Anterior Cingulate Cortex Following Digit Amputation in the Anaesthetized Rat.' *Journal of Physiology* 532: 823.

- Weisenberg, M., Tepper, I., & Schwarzwald, J. (1995) 'Humor as a Cognitive Technique for Increasing Pain Tolerance.' *Pain* 63: 207–212.
- Weiskrantz, L. (1996) 'Blindsight Revisited.' *Current Opinion in Neurobiology* 6: 215–220.
- Wiggins, D. (1987) 'A Sensible Subjectivism?' In *Needs, Values, Truth*, Oxford: Blackwell Publishers.
- Williams, A.C. (2003) 'Facial Expression of Pain: An Evolutionary Account.' *The Behavioral and Brain Sciences* 25: 439–488.
- Williams, B. (1970) 'The Self and the Future.' *The Philosophical Review* 79(2): 161–180.
- Williams, B. (1995) 'Internal and External Reasons.' In *Moral Luck*, Oxford: Oxford University Press.
- Woolf, C.J. (2011) 'Central Sensitization: Implications for the Diagnosis and Treatment of Pain.' *Pain* 152: S2–S15.
- Wundt, W. (1897/1980) *Outlines of Psychology*. Springer.
- Xu, H., Wu, L., Wang, H., Zhang, X., Vadakkan, K., Kim, S., Steenland, H., & Zhuo, W. (2008) 'Presynaptic and Postsynaptic Amplifications of Neuropathic Pain in the Anterior Cingulate Cortex.' *Neurobiology of Disease* 28(29): 7445–7453.
- Yamamoto, S. & Kitazawa, S. (2001a) 'Reversal of Subjective Temporal Order Due to Arm Crossing.' *Nature Neuroscience* 4(7): 759–765.
- Yamamoto, S. & Kitazawa, S. (2001b) 'Sensation at the Tips of Invisible Tools.' *Nature Neuroscience* 4(10): 979–980.
- Yancey, P. & Brand, P. (1997) *The Gift of Pain*. Grand Rapids, MI: Zondervan Publishing House.
- Young, P.T. (1959) 'The Role of Affective Processes in Learning and Motivation.' *Psychological Review* 66: 104–125.
- Younger, J.W., Chu, L.F., D'Arcy, N.T., Trott, K.E., Jastrzab, L.E., & Mackey, S.C. (2011) 'Prescription Opioid Analgesics Rapidly Change the Human Brain.' *Pain* 152: 1803–1810.
- Zillmann, D., Rockwell, S., Schweitzer, K., & Sundar, S.S. (1993) 'Does Humor Facilitate Coping with Physical Discomfort?' *Motivation & Emotion* 17(1): 1–21.
- Zubieta, J.-K., Smith, Y.R., Bueller, J.A., Xu, Y., Kilbourn, M.R., Jewett, D.M., Meyer, C.R., Koeppe, R.A., & Stohler, C.S. (2001) 'Regional Mu Opioid Receptor Regulation of Sensory and Affective Dimensions of Pain.' *Science* 293: 311–315.
- Zubieta, J.-K., Bueller, J.A., Jackson, L.R., Scott, D.J., Xu, Y., Koeppe, R.A., Nichols, T.E., & Stohler, C.S. (2005) 'Placebo Effects Mediated by Endogenous Opioid Activity on Mu-Opioid Receptors.' *The Journal of Neuroscience* 25: 7754–7762.

Index

- Affect 2–7, 27–59, 61, 64–73, 77, 93, 95, 147, 151, 153, 155–7, 161, 169, 171, 181, 190–1; as negative valence 127–32; *see also* Pain, affective (motivational) component of
- Affordances 187
- Anterior cingulate cortex (ACC) 122, 131, 136, 138, 153–5, 157, 159–60, 175
- Aydede, Murat 2–3, 27–59
- Bain, David 1–9, 33, 38, 43, 55–6, 58, 91, 158, 178–9, 181
- Behaviour 7, 88–90, 182, 186, 191
- Belief 3, 31, 36, 38–46, 48–52, 54–8, 114–8
- Benedetti, Fabrizio 96–7, 101, 108, 110
- Body map 189–90
- Bodily ownership (sense of) 8, 179–92
- Boly, M. 173
- Brady, Michael 1–9, 81–94
- Care 57–8, 170, 181, 186, 188, 190
- Commands 2–3, 13–9, 21–5, 30–3, 38, 43, 46–52, 55, 57–8, 114, 128–30, 132–5, 139, 141, 150, 181–2; *see also* Imperatives
- Conditioning 64, 73–4, 96, 98, 134, 157, 160
- Consciousness 19, 35, 54, 73, 149, 151, 159; disorders of 7–8, 163–77; moral significance of 167–70
- Content 2, 5, 7, 13–20, 22–5, 28–36, 39, 41–3, 45–6, 48–58, 91, 127, 149–50
- Corns, Jennifer 1–9, 55, 95–119
- Cutter, Brian 15–6, 25, 91
- D'Arms, Justin 83
- Desire 24, 31, 35–6, 38–58, 62, 67, 93, 101, 116, 149–51, 156, 161, 168–9
- Desire-based accounts of pain/unpleasantness 2–3, 129
- de Vignemont, Frédérique 8, 178–92
- Dokic, Jérôme 178–9
- Dopamine/dopaminergic signalling 3, 65–7
- Ellingsen, Dan-Mikael 3, 60–77
- Emotion(s) 1, 4–6, 31, 43–4, 61, 65, 76–7, 81–9, 91–3, 95–6, 115, 120, 122–33, 135–8, 140–1, 143, 153, 163, 169, 176
- Evaluation 35, 53, 85–6, 88, 90–2, 125
- Evaluativism 2, 29, 33–4, 37–40, 42–4, 48, 52, 55–6, 81, 91–2
- Facial expression (of pleasure, pain, emotion) 3, 60, 63–4, 123–5
- First-person 8, 142, 178–92
- fMRI activation/studies 64–5, 67, 74–5, 121–2, 136–7, 152, 163–4, 166
- Fulkerson, Matthew 2–3, 27–59
- Functionalism 28, 34–5, 57
- Grahek, Nikola 155
- Greene, C. S. 109
- Hall, Richard 13, 25, 31–3, 46–7
- Hall, Wayne 18
- Hardcastle, Valerie Gray 7–8, 163–77
- Hedomotive 38
- Hedonic tone 28, 95–7
- Hedonic value 55, 60–4, 67, 69–73, 76–7
- Imperatives 13–25, 31, 48, 50, 52, 132
- Imperativism 2–3, 13–26, 29–33, 37–9, 46–51, 55, 58
- Injury 6, 17, 38, 81, 87–8, 176, 181, 187, 191

- Intensity (of commands, liking, pain)
2, 13–26, 62, 66, 71–2, 98, 100, 124,
148–9, 161–2, 175, 183–4, 188
- Intentional content 2, 7, 14, 31,
48, 149–50
- Intentionalism 16, 18–9, 24, 30–1, 47–8;
see also Representation(alism)
- International Association for the Study of
Pain (IASP) 4, 123, 163, 170
- Jacobson, Daniel 83, 92
- James, William 86, 126–7
- Justification 39–45, 48–51, 55–8
- Klein, Colin 2, 13–26, 31–2, 39, 47, 53,
57–8, 158, 181–2
- Korsgaard, Christine 150–1, 161
- Kringelbach, Morten L. 3, 60–77
- Leknes, Siri 3, 60–77
- Liking/disliking 2–4, 28, 53, 62–4, 66–7,
69, 75, 77, 139, 141, 151
- Locked-In Syndrome 7, 164
- Maravita, A. 184
- Martínez, Manolo 13–26
- Masochism 77
- McGill Pain Questionnaire 121, 140
- Melzack, R. 120–1, 123, 131, 147, 161
- Motivation 1–4, 6–7, 17, 35–6, 38–40,
43–52, 55–8, 61–4, 68, 71, 74, 77, 123,
128–30, 134, 139, 141, 147, 149–50,
156, 158–61, 171, 175, 180–2, 191
- Naturalism 19, 34, 51–2, 54–7
- Nociception/nociceptive processing 74,
170–5, 179
- Nonhuman animals 6–7, 62, 64–7, 72, 77,
147, 151–60
- Nunn, R. 111
- Pain: affective (motivational) component
of 3, 5–6, 120–43, 147, 150, 157, 159–
61; affective modulation of 73; alien 8,
179, 182–5, 191–2; asymbolia 1, 3, 27,
91, 155–8, 191; badness of 2–3, 6, 42–4,
46, 52, 55–6, 58, 121, 138–40, 143,
161, 176; chronic 1, 8–9, 61, 81, 122,
125, 142, 152, 174–6; delusions 8, 183,
188–9, 192; deviant 1, 6–8; dissociation
between affective and sensory
components 6–7, 121, 147–50, 155–6;
felt quality of 28–9, 39; intensity of 2,
13–26, 62, 71–2, 98, 124, 148–9, 161–2,
175, 183–4, 188; location of 25, 32, 46,
56, 121, 180–1, 190, 192; malleability of
68–9; neuroscience of 3, 34–5, 60–77,
136, 153; pathways 7, 74, 77, 97, 104,
121, 155, 160, 165, 170–1; phantom
limb 180–1, 183, 192; processing 2–3,
7–8, 35–6, 38, 44–7, 51–2, 54, 56, 60–3,
65, 67, 69, 72–5, 77, 121, 132, 138,
153–4, 161, 167, 170–6, 184; sensitivity
121, 132, 137–8, 186, 190; sensory
(discriminative) component of 6–7, 51,
58, 64, 68, 121–2, 127, 129–30, 134,
136, 139, 142, 147–8, 161, 171
- Parfit, Derek 150–1, 161
- Pautz, Adam 17, 25–6
- Persistent vegetative states 164–5,
170, 173
- Phenomenology 29–30, 32–3, 35–6, 45–6,
51, 53–4, 56–7, 86, 91, 130, 138–43,
150, 161
- Pia, Lorenzo 178–9
- Placebo effect(s) 5, 73–5, 77, 95–119
- Pleasure 3–4, 28–9, 43, 53–4, 56, 60–77,
122, 124, 129–30, 167–8
- Prinz, Jesse 4–6, 120–43
- Psychofunctionalism 29, 34–8, 44–7,
51, 54–6
- Rationality 4, 49, 81–94, 168
- Reason(s) 3–5, 14, 27–59, 81, 83–5, 87,
91–3, 181
- Representation(alism) 3, 7, 13–4, 26,
27–34, 38–40, 42–3, 46–8, 51–5, 57–8,
68, 73, 89, 130, 149, 151, 188–90
- Rubber Hand Illusion 8, 182–5, 188, 192
- Schnakers, C. 173–4
- Sensations/sensory experiences 8, 13, 17,
25–6, 27–9, 32, 35, 41, 43–4, 51, 53, 55,
58, 60–1, 64–5, 68, 75–7, 90, 128–30,
133, 142, 148–51, 161, 175, 179–82,
184–6, 189–91
- Sensory affect *see* Affect
- Shriver, Adam 6–7, 147–62
- Spinoza, Baruch 4, 122
- Suffering 1, 4–5, 81–94, 131, 149, 151,
170, 175
- Threat(s) 4, 8, 13, 45, 50, 57, 60, 77, 90,
126, 156, 179, 183, 185, 187–9
- Tye, Michael 15–6, 25, 91
- Unpleasantness 1–4, 6–7, 27–9, 32–40,
43–6, 48, 51, 55, 60, 62, 64, 69–70, 72,

- 74–5, 77, 91, 95–7, 100, 102, 121–3,
126, 138–43, 147–62, 163, 170, 180–1,
191; *see also* Pain, badness of
- Valence 6, 27–8, 35, 61, 120, 127–43,
181, 190
- Value(s) 5, 23, 28–9, 33, 46, 53, 55, 57,
60–4, 67, 69–73, 76–7, 82–93, 130, 148,
150–1, 158, 161, 167–8, 171, 190
- Wanting 2–4, 28, 62, 65–7; *see also* Desire;
Liking/disliking